

Best Local Similarity 34.2%; Pred. No. 1.2e-19;  
Matches 64; Conservative 48; Mismatches 64; Indels 11; Gaps 3;

QY 10 INFETAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPINQRFILVELSNHAEL 69  
DB 3 VSFSLGADSKSYSFITRALRKALPSKEKVSNIPLLPASGA---SYILMQLSNYDAK 59  
QY 70 SVTLALDVTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQVRYTFAFGNVDRLQ 129  
DB 60 AITMAIDVTNAYVIMGYLVNSTSYF---ANESDAKLASQVYFKGSTLVITIPYSGNYERLQN 116  
QY 130 LAGNRENIELGNGPLEEALSYLYSTGGTQPLTARSFIICIQMISEAARFQVIEGEM 189  
DB 117 AAGKIREKIPGLFRALDSALTSIFHYDS-----TAAAFVLVILQTTAASRKYIEGOI 171  
QY 190 RTRIRYN 196  
DB 172 IERIPKN 178

## RESULT 14

S62627  
agglutinin I precursor - European elder  
C:Species: Sambucus nigra (European elder)  
C:Date: 24-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 20-Aug-1999  
C:Accession: S62627; S62619  
R:van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.  
Eur. J. Biochem. 235, 128-137, 1996  
A:Title: The NeuA(alpha-2,6)-Gal/GalNAc-binding lectin from elderberry (Sambucus nigra)  
A:Reference number: S62619, MUID:96202926; PMID:8631319  
A:Accession: S62627  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-570 <VAN>  
A:Cross-references: EMBL:U27122; NID:g1141772; PIDN:AAC49158.1; PID:g1141773  
A:Accession: S62619  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 29-39;309-319 <VA2>  
C:Superfamily: ricin; rRNA N-glycosidase homology  
F:37-283/Domain: rRNA N-glycosidase homology <RNG>

Query Match 28.3%; Score 291; DB 2; Length 570;  
Best Local Similarity 39.0%; Pred. No. 2.4e-18;  
Matches 76; Conservative 33; Mismatches 74; Indels 12; Gaps 6;  
QY 2 IFPKQYPIINFETAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPINQRFILV 61  
DB 29 VTPPVYVSVSFNLTA--DYEPFLALQEKVILGNHTAFDLPVLPNPSQVSDSNRFVLV 86  
QY 62 ELSNHAELSVTLALDVTNAYVGVYRAGNSAYFFHPDN--QEDAEATHLFTDVQVRYTFA 119  
DB 87 PLTNPSGDVTTLAIDVNVLYVAFSSNGKSYFFSGSTAVQD-----NLFVDT-TOEELN 140  
QY 120 FCGNVDRLQLAGNRENIELGNGPLEEALSYLYS--TGGTQPLTARSFIICIQMISE 178  
DB 141 FTGNYTSLEKQVGRVYIPLPKSLDQALSSRTYTLTAGDTKP-LARGLLVVIQWSE 199  
QY 179 AARFQVIEGEMTRI 193  
DB 200 AARFRIELRIRTSI 214

## RESULT 15

PD0018  
mistletoe lectin I A chain - Viscum album (fragment)  
C:Species: Viscum album  
C:Date: 03-Jul-1998 #sequence\_revision 10-Jul-1998 #text\_change 26-Aug-1999  
C:Accession: PD0018  
R:Eschenburg, S.; Krauspenhaar, R.; Mikhailov, A.; Stoeva, S.; Betzel, C.; Voelter, W.  
Biochem. Biophys. Res. Commun. 247, 367-372, 1998  
A:Title: Primary structure and molecular modeling of mistletoe lectin I from Viscum albu  
A:Reference number: PD0018; MUID:98308123; PMID:9642133

A:Accession: PD0018

A:Molecule type: protein

A:Residues: 1-254 <ESC>

C:Superfamily: ricin; rRNA N-glycosidase homology

F:7-246/Domain: rRNA N-glycosidase homology <RNG>

Query Match 27.8%; Score 286; DB 2; Length 254;

Best Local Similarity 38.4%; Pred. No. 2.4e-18;

Matches 81; Conservative 36; Mismatches 56; Indels 38; Gaps 10;

QY 14 TAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPIN--QRFILVELSNHAELSV 71  
DB 9 THQTTGEEYFRFITLLRDYVSSGS-FSNEIPLL-RQSTIPVSDAQRFLVELTNGQDSV 66  
QY 72 TLALDVTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQVRYTFAFGNVDRLQLA 131  
DB 67 TAAIDVTNAYVAVYQAGDQSYFLR-DAPRGAE--THLFTGT-TRSSLPFNGSYPOLERYA 122  
QY 132 GNLRNIELGNGPLEEALSYLYSTGGTQPLTARSFIICIQMISEAARF-----Q 183  
DB 123 GH-RDQIPFLGIDQLIQSVTALRF---PGSSTFQARSILILLIOMISEAARFNPILWRYRQ 178  
QY 184 YIE-----GEMTRIRYN 196  
DB 179 YINSGASFLPDVYMLELETSMGQQSTQVQHS 209

Search completed: February 10, 2004, 16:28:03

Job time : 10.8228 secs





```
CC -----
DR EMBL; M12089; AAA33869.1; -.
DR EMBL; S40368; AAB22584.1; -.

```

AC	P28590;
DT	01-DEC-1982 (Rel. 24, Created)
DT	01-DEC-1982 (Rel. 24, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	ABRIN-C precursor [Contains: ABRIN-C A chain
DE	(EC 3.2.2.22); ABRIN-C B chain].
OS	Abrus precatorius (Indian licorice) (Crab's e

```

Best Local Similarity   45.5%;    Pred. NO. 2.6e-25;
Matches      86;     Conservative      24;    Mismatches       70;    Indels        9;    Gaps          4;

QY      6 QYPINFTTAGATVQSYTFNFIKAVRGRLTIGADVRIEIPVLKNRVGLP.INQREITLVELSN 65
Db      35 QDOVKIFTEGATSQSOKFIEALRQLRTGG--LIHDIPLVPDPTTVVEENRYITVELSN 92
QY      66 HAELSVTLDIVNAVVGVRAGNSAYFFHPDNQDEAEAILHTLFTDVQNRYTFAFGNYD 125
Db      93 SERESIEVGIDIVNAVVAIRAGSQSYEL---RDAPASASTYLFPQTG-QYSLEFDGSYG 148
QY      126 RLEQLAGNLRENIELNGPLERBAISALYYSTGTOLPTLARSGIICIMISEARFOYI 185
Db      149 DLERWAHQTREEISLGILQAALTHAIS----FLRSGASNDEEKARTLLVIIONMASEAARYEI 205
QY      186 EGEGRTRIR 194
Db      206 SNRWGVSIIR 214

RESULT 4
ABRB_ABRPR
ID -ABRB_ABRPR STANDARD; PRT; 527 AA.
AC Q06077; F81374;
DT 15-DEC-1998 (Rel. 37, Created)
DD 15-DEC-1998 (Rel. 37, Last sequence update)
DR 28-FEB-2003 (Rel. 41, Last annotation update)
DE Abrin-b precursor [Contains: Abrin-b A chain (rRNA N-glycosidase)
(EC 3.2.2.2); Abrin-b B chain].
DS Abrus precatorius (Indian licorice) (Crab's eye).
OS Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
CX NCBI_TaxID=3816;
RX SEQUENCE FROM N.A.
EP MEDLINE=93132798; PubMed=8421313;
RA Hung C.-H., Lee M.-C., Lee T.C.; Lin J.-Y.;
RT "Primary structure of three distinct isoforms determined by cDNA
sequencing. Conservation and significance.";
RL J. Mol. Biol. 229:263-267(1993).
[2]
SEQUENCE OF 260-527.
RP TISSUE=Seed;
RC MEDLINE=93169023; PubMed=7763422;
RX Kimura M., Sumizawa T., Funatsu G.;
RA "The complete amino acid sequences of the B-chains of abrin-a and
abrin-b, toxic proteins from the seeds of Abrus precatorius.";
RL Biosci. Biotechnol. Biochem. 57:166-169(1993).
CC -! FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
SUBUNIT BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
CC ABRAIN-A IS MORE TOXIC THAN GALICIN.
CC -! FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
PRECEDES ENDOCYTOSIS.
CC -! CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
specific adenosine on the 28S rRNA.
CC -! SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC -! DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC -! SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
CC -! SIMILARITY: Contains 2 ricin B-type lectin domains.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on it
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercia
entities requires a license agreement (See http://www.isb-sib.ch/announce
or send an email to license@isb-sib.ch).
```

34.4%; Score 353.5; DB 1; Length 562;

```

DR EMBL; M98345; AAA32625.1; -.
DR PIR; S32430; S32430.
DR HSSP; P41140; LABR.
DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS0231; RICIN B LECTIN; 2.
DR PROSITE; PS0275; SHIGA-RICIN; 1.
KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
KW Glycoprotein; Lectin; Pyrrolidone carboxylic acid.
FT CHAIN 1 250
FT PEPTIDE 251 260
FT CHAIN 261 527
FT DOMAIN 272 399
FT DOMAIN 402 526
FT REPEAT 402 324
FT REPEAT 325 365
FT REPEAT 368 400
FT REPEAT 413 448
FT REPEAT 452 491
FT REPEAT 494 527
FT ACT_SITE 163
FT DISULFID 246
FT DISULFID 285
FT DISULFID 328
FT DISULFID 416
FT DISULFID 455
FT MOD_RES 1 1
FT CARBOHYD 110 110
FT CARBOHYD 360 360
FT CARBOHYD 400 400
FT CONFLICT 282 282
FT CONFLICT 291 291
FT CONFLICT 350 351
FT CONFLICT 378 378
FT CONFLICT 426 426
FT CONFLICT 428 428
FT CONFLICT 431 431
FT CONFLICT 484 484
FT CONFLICT 491 491
FT CONFLICT 493 493
FT CONFLICT 502 502
FT CONFLICT 509 509
FT CONFLICT 513 513
FT CONFLICT 516 516
FT SEQUENCE 527 AA; 59114 MW; 3253AE490CE9494A CRC64;
Query Match 33.5%; Score 345; DB 1; Length 527;
Best Local Similarity 45.5%; Pred. No. 1.5e-24;
Matches 87; Conservative 22; Mismatches 72; Indels 10; Gaps 4;
QY 6 QYPIINFTAGATVOSYNNFTRAVRGRITGADVREHPVLEPNRVGLPINQRFILVELSN 65
Db 1 QDQVVKFTTEGATSSQYKQFIEALRQLTGG--LIHGIPVLPDPTTLQERNRYISVELSN 58
QY 66 HAEISVTLALDVYNAVYVGYRAGNSAYRFPHPDQNEADAEITLFTDVQNRVYTFAGGNYD 125
Db 59 STOTESIAGIDVSNVAVYRAGNSYFL---RDAPTSASRYLFTGTQ-QYSLRNGSYI 114
QY 126 RLEQAGNLRENIELNGPLEBAISALYYVYTGQTLPFLARSPFICIMISEARFQYI 185
Db 115 DLERUARQTRQIPLGLQALRAISFL---QSGTDDQEIARTLIVLIQWASEAARYRPI 170
QY 186 EGEIMTRIRYN 196
Db 171 SYRVGVGSIRTN 181
RESULT 5

```

```

RIPT_TRIKI
ID RIPT_TRIKI STANDARD; PRT; 289 AA.
AC P09889;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ribosome-inactivating protein alpha-trichosanthin precursor
DE (trRNA N-glycosidase) (EC 3.2.2.22) (Alpha-TCS).
OS Trichosanthin kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID:3677;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Maximowicz;
RX MEDLINE=91153657; PubMed=1999291;
RA Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;
RT "Cloning of trichosanthin cDNA and its expression in Escherichia
RL coli."; Gene 97:267-272(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Maximowicz; TISSUE=Leaf;
RX MEDLINE=90256790; PubMed=2341400;
RA Chow T., Feldman R.A., Lovett M., Piatak M.;
RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a
RL type I ribosome-inactivating protein.";
RN [3]
RP SEQUENCE OF 24-270.
RC STRAIN=Maximowicz; TISSUE=Tuberos root;
RX MEDLINE=90256789; PubMed=2341399;
RA Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R.,
Wu P., Hwang K., Piatak M.;
RT "Primary amino acid sequence of alpha-trichosanthin and molecular
RL models for abrin A-chain and alpha-trichosanthin.";
RN [4]
RP SEQUENCE OF 24-270.
RC TISSUE=Tuberos root;
RX Wang Y., Qian R.Q., Zhang L.Q., Xia Z.X.,
Tian G.Y., Ni C.Z.;
RT "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and
RL application."; Pure Appl. Chem. 58:789-798(1986).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).
RX MEDLINE=94344957; PubMed=8066085;
RA Zhou F., Fu Z., Chen M., Lin Y., Pan K.;
RT "Structure of trichosanthin at 1.88-A resolution.";
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX MEDLINE=95344383; PubMed=7619070;
RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
RT "Studies on crystal structures, active centre geometry and
RL depurinating mechanism of two ribosome-inactivating proteins.";
Biochem. J. 309:285-298(1995).
CC -!- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS
CC CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT
CC INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

```
CC CC -----
DR DR EMBL; M34858; AAA34207.1; -.
DR DR EMBL; J05434; AAA34206.1; -.
DR DR PIR; J0566; RLTTZT.
DR DR PDB; 1MRJ; 07-FEB-95.
DR DR PDB; 1MRK; 07-FEB-95.
DR DR PDB; 1TCS; 10-JUL-95.
DR DR PDB; 1J4G; 28-JAN-03.
DR DR PDB; 1NLI; 21-JAN-03.
DR DR PDB; 1QD2; 24-APR-00.
DR DR InterPro; IPR001574; RIP.
DR DR Pfam; PF00161; RIP: 1.
DR DR PRINTS; PR00396; SHIGARICIN.
DR DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin; Signal; 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 270
FT FT
FT FT RIBOSOME-INACTIVATING PROTEIN ALPHA-
FT FT TRICHOSANTHIN.
FT FT MISSING IN MATURE PROTEIN.
FT FT BY SIMILARITY.
FT FT IPDL -> LPLI (IN REF. 4).
FT FT MISSING (IN REF. 4).
FT FT I -> L (IN REF. 4).
FT FT V -> VDAGLPRNAVL (IN REF. 4).
FT FT KI -> GL (IN REF. 4).
FT FT K -> S (IN REF. 4).
FT FT WS -> LWL (IN REF. 4).
FT FT Q -> T (IN REF. 4).
FT FT S -> T (IN REF. 2).
FT FT MISSING (IN REF. 4).
FT FT T -> M (IN REF. 2).
FT FT
FT FT STRAND 25 28
FT FT TURN 30 31
FT FT
FT FT HELIX 34 46
FT FT TURN 47 47
FT FT STRAND 50 54
FT FT TURN 55 56
FT FT STRAND 57 60
FT FT HELIX 66 69
FT FT STRAND 70 76
FT FT TURN 78 79
FT FT STRAND 82 88
FT FT TURN 89 92
FT FT STRAND 93 99
FT FT TURN 100 101
FT FT STRAND 102 105
FT FT HELIX 109 114
FT FT TURN 115 117
FT FT TURN 120 121
FT FT STRAND 124 127
FT FT HELIX 134 141
FT FT TURN 142 142
FT FT HELIX 145 147
FT FT STRAND 150 150
FT FT HELIX 152 163
FT FT TURN 164 165
FT FT HELIX 167 180
FT FT TURN 181 181
FT FT HELIX 182 186
FT FT STRAND 187 187
FT FT HELIX 188 195
FT FT TURN 196 196
FT FT STRAND 202 202
FT FT HELIX 206 226
FT FT TURN 227 230
FT FT STRAND 231 239
FT FT TURN 241 242
FT FT STRAND 245 250
FT FT TURN 251 252
FT FT HELIX 254 258
FT FT TURN 259 259
CC CC -----
or send an email to license@isb-sib.ch.
```

FT STRAND 260 260  
FT STRAND 263 263  
FT TURN 266 268  
SQ SEQUENCE 289 AA; 31676 MW; 5CE09BB630575BB9 CRC64;  
Query Match 33.2%; Score 342; DB 1; Length 289;  
Best Local Similarity 39.5%; Pred. No. 1.4e-24;  
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;  
QY 10 INFATTAGATVQSYNTFIRAVRGRLTTGADVRHEIPVLPNRVGLPINORFILLVLSNHAEL 69  
DB 25 VSFRLSGATSSSYGVFISNRKALPNERKL-YDIPLL--RSSLPGSRYSALHILHTINVADE 81  
QY 70 SVTTLALDVTNAYVVGVRAGNSAYFFPHDPNOEDA-EAITHLFTDVQNRXYTFAFGNYDRLE 128  
DB 82 TISVAIDVTNVIYINGRAGDTSYFF--NEASATEAAKYVFKDAMEKVKVLPYSGNYERLQ 138  
QY 129 QLAGNLRNTELGNGPLEBAISALYYVYSTGTQLPILARSFIFCIOMISEARFOVIEGE 188  
DB 139 TAAGKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSERARYKFEQQ 193  
QY 189 METRI 193  
DB 194 IGRV 198  
RESULT 6  
RIP2.BRYDI  
ID RIP2.BRYDI STANDARD; PRT; 282 AA.  
AC P98184; Q998J0;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ribosome-inactivating protein bryodin II precursor (rRNA N-glycosidase) (EC 3.2.2.22) (BD2).  
OS Bryonia dioica (Red bryony).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.  
OX NCBI\_TaxID=3652;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Siegall C.B., Gawlak S.L., Marquardt H.;  
RT "Bryodin 2 a ribosome-inactivating protein isolated from the plant Bryonia dioica.";  
RL Patent number US5597569, 28-JAN-1997.  
RN [2]  
RP SEQUENCE OF 22-42.  
RC TISSUE=Root;  
RX MEDLINE=95151812; PubMed=7849072;  
RA Siegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B., Marquardt H.;  
RT "Characterization of ribosome-inactivating proteins isolated from Bryonia dioica and their utility as carcinoma-reactive immunoconjugates.";  
RL Bioconj. Chem. 5:423-429(1994).  
CC -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 1 RIP SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; I34238; -; NOT ANNOTATED\_CDS.  
CC HSP; P09989; 1MRJ.  
DR DR





CC PROTEIN SYNTHESIS BUT DOES NOT AFFECT PLANT NOR BACTERIAL PROTEIN  
 CC SYNTHESIS. THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE  
 CC B-CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE  
 CC BINDING OF NIGRIN B TO THE CELL MEMBRANE THAT PRECEDES  
 CC ENDOCYTOSIS.  
 CC  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-  
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U41299; AAB39475.1; --  
 CC PIR: S37382; S37382; Ricin\_B\_lectin.  
 CC InterPro: IPR000772; Ricin\_B\_lectin.  
 CC InterPro: IPR001574; RIP.  
 CC Pfam: PF00652; Ricin\_B\_lectin; 6.  
 CC Pfam: PF00161; RIP; 1.  
 CC PRINTS: PR00396; SHIGARICIN.  
 CC SMART: SM00458; RICIN; 2.  
 CC PROSITE: PS00275; SHIGA\_RICIN; 1.  
 CC PROSITE: PS0231; RICIN\_B\_LECTIN; 2.  
 CC K plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 CC Glycoprotein; Lectin; Signal.  
 CC SIGNAL 1 25  
 CC CHAIN 26 297 NIGRIN B A CHAIN.  
 CC CHAIN 298 563 NIGRIN B B CHAIN.  
 CC DOMAIN 305 431 RICIN B-TYPE LECTIN 1.  
 CC DOMAIN 434 559 RICIN B-TYPE LECTIN 2.  
 CC REPEAT 316 356 1-ALPHA.  
 CC REPEAT 357 397 1-BETA.  
 CC REPEAT 400 432 1-GAMMA.  
 CC REPEAT 445 482 2-ALPHA.  
 CC REPEAT 486 524 2-BETA.  
 CC REPEAT 527 554 2-GAMMA.  
 CC ACT SITE 188 188 BY SIMILARITY.  
 CC DISULFID 274 302 INTERCHAIN (BY SIMILARITY).  
 CC DISULFID 319 338 BY SIMILARITY.  
 CC DISULFID 360 377 BY SIMILARITY.  
 CC DISULFID 448 463 BY SIMILARITY.  
 CC DISULFID 489 506 BY SIMILARITY.  
 CC CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CONFLICT 39 39 K -> V (IN REF. 2).  
 CC SEQUENCE 563 AA; 62300 MW; F250CBE24621BF14 CRC64;  
 CC  
 CC Query Match 32.1%; Score 330.5; DB 1; Length 563;  
 CC Best Local Similarity 39.5%; Pred. No. 3.7e-23;  
 CC Matches 75; Conservative 39; Mismatches 61; Indels 15; Gaps 4;  
 CC  
 CC 7 YPIINFITAGATVSYTNFIRAVRGLTGDVREHPIVLPNVRGLPINORFILVELSNH 66  
 CC 28 YPSVSNLIDGAKSYRDFLSNLRKTKVATGTYEVNGLPVLRRSEVQKSRFVLPLTNY 87  
 CC 67 AELSVTLALDYNVAVGVYRAGNAGYFPHPDQDAEAI--THLFTDQVNEYTFAGGNY 124  
 CC 88 NGNTVTAVDVTNLYVAFSGNANSYFF-----KDTEVQKSNLFGVTKQN-TLSFTGNY 141  
 CC 125 DRLEQLAGNLRNLELGNPLLEEISALYYSTGTQTLPLTLAGFIICMISEAARFOY 184

Db 142 DNLEAANTRRSEIEGLSPIDGNAITSLYHGD-----SVARSLLVVIQWSEARFRY 194  
 Qy 185 IEGEMRTRIR 194  
 Db 195 IEQEVRRSLQ 204  
 RESULT 9  
 ID ABRA ABRPR STANDARD; PRT; 528 AA.  
 AC P11140; P28589;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Abrin-a precursor [Contains: Abrin-a A chain (rRNA N-glycosidase)  
 (EC 3.2.2.22); Abrin-a B chain].  
 OS Abrus precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 OC NCBI\_TaxID=3816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93132798; PubMed=8421313;  
 RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;  
 RT "Primary structure of three distinct isoabrinins determined by cDNA  
 sequencing. Conservation and significance."  
 RL J. Mol. Biol. 229:263-267(1993).  
 RN [2]  
 RP SEQUENCE OF 1-251.  
 RC TISSUE=Seed;  
 RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;  
 RT "The complete amino acid sequence of the A-chain of abrin-a, a toxic  
 protein from the seeds of Abrus precatorius."  
 RL Agric. Biol. Chem. 52:1095-1097(1988).  
 RN [3]  
 RP SEQUENCE OF 1-251 FROM N.A.  
 RC TISSUE=Leaf;  
 RX MEDLINE=91201329; PubMed=2016300;  
 RA Evensen G., Mathiesen A., Sundan A.;  
 RT "Direct molecular cloning and expression of two distinct abrin  
 A-chains."  
 RL J. Biol. Chem. 266:6848-6852(1991).  
 RN [4]  
 RP SEQUENCE OF 262-528.  
 RX MEDLINE=92371656; PubMed=1505674;  
 RA Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;  
 RT "The complete primary structure of abrin-a B chain."  
 RL FEBS Lett. 309:115-118(1992).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).  
 RX MEDLINE=95333188; PubMed=7608980;  
 RA Tahirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;  
 RT "Crystal structure of abrin-a at 2.14 A."  
 RL J. Mol. Biol. 250:354-367(1995).  
 CC -!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.  
 CC ABRIN-A IS MORE TOXIC THAN RICIN.  
 CC -!- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
 CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
 CC PRECEDES ENDOCYTOSIS.  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -!- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-  
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; M98344; AAA32624.1; ALT\_INIT.  
 DR EMBL; X54872; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; S32429; TZLSA.  
 DR PDB; 1ABR; 07-FEB-95.  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 KW Glycoprotein; Lectin; 3D-structure; Pyrrolidone carboxylic acid.  
 FT CHAIN 1 251  
 FT PEPTIDE 252 261  
 FT CHAIN 262 528  
 FT DOMAIN 273 400  
 FT DOMAIN 403 527  
 FT REPEAT 326 366  
 FT REPEAT 369 401  
 FT REPEAT 414 449  
 FT REPEAT 453 492  
 FT REPEAT 495 528  
 FT ACT\_SITE 164 164  
 FT DISULFID 247 269  
 FT DISULFID 286 305  
 FT DISULFID 329 346  
 FT DISULFID 417 430  
 FT DISULFID 456 473  
 FT MOD\_RES 1 1  
 FT CARBOHYD 361 361  
 FT CARBOHYD 401 401  
 FT CONFLICT 202 202  
 FT CONFLICT 298 298  
 FT CONFLICT 427 427  
 FT CONFLICT 467 467  
 FT CONFLICT 483 483  
 FT STRAND 5 8  
 FT TURN 10 11  
 FT TURN 14 28  
 FT STRAND 32 33  
 FT TURN 34 35  
 FT TURN 42 43  
 FT STRAND 47 49  
 FT STRAND 51 57  
 FT STRAND 63 69  
 FT TURN 70 72  
 FT STRAND 75 79  
 FT STRAND 83 86  
 FT TURN 88 89  
 FT TURN 92 93  
 FT TURN 94 97  
 FT TURN 100 101  
 FT STRAND 103 106  
 FT TURN 113 114  
 FT TURN 115 119  
 FT TURN 124 126  
 FT STRAND 129 129  
 FT HELIX 131 142  
 FT TURN 143 144  
 FT HELIX 148 167  
 FT STRAND 168 168  
 FT HELIX 169 180  
 FT TURN 181 182

FT STRAND 185  
 FT HELIX 186  
 FT TURN 187  
 FT HELIX 197  
 FT TURN 198  
 FT STRAND 207  
 FT TURN 212  
 FT TURN 220  
 FT STRAND 222  
 FT STRAND 223  
 FT TURN 226  
 FT TURN 232  
 FT HELIX 233  
 FT STRAND 235  
 FT STRAND 240  
 FT STRAND 243  
 FT STRAND 248  
 FT STRAND 268  
 FT STRAND 276  
 FT STRAND 277  
 FT STRAND 279  
 FT STRAND 282  
 FT STRAND 286  
 FT STRAND 289  
 FT HELIX 290  
 FT TURN 292  
 FT TURN 296  
 FT STRAND 299  
 FT HELIX 311  
 FT STRAND 313  
 FT TURN 315  
 FT TURN 319  
 FT STRAND 320  
 FT STRAND 322  
 FT TURN 325  
 FT TURN 326  
 FT STRAND 327  
 FT TURN 328  
 FT TURN 337  
 FT STRAND 338  
 FT TURN 340  
 FT TURN 346  
 FT HELIX 351  
 FT STRAND 353  
 FT STRAND 355  
 FT STRAND 357  
 FT TURN 361  
 FT STRAND 364  
 FT TURN 367  
 FT STRAND 370  
 FT TURN 371  
 FT TURN 374  
 FT STRAND 380  
 FT STRAND 381  
 FT STRAND 383  
 FT STRAND 385  
 FT HELIX 393  
 FT STRAND 395  
 FT STRAND 398  
 FT STRAND 406  
 FT STRAND 410  
 FT STRAND 411  
 FT HELIX 413  
 FT STRAND 417  
 FT TURN 421  
 FT TURN 422  
 FT STRAND 423  
 FT TURN 424  
 FT TURN 428  
 FT TURN 432  
 FT HELIX 433  
 FT STRAND 435  
 FT STRAND 439  
 FT TURN 441  
 FT STRAND 443  
 FT STRAND 444  
 FT TURN 447  
 FT TURN 449  
 FT TURN 451  
 FT STRAND 452  
 FT TURN 459  
 FT STRAND 464  
 FT STRAND 465  
 FT STRAND 467  
 FT TURN 472  
 FT TURN 474  
 FT STRAND 475  
 FT HELIX 478  
 FT STRAND 480  
 FT TURN 484  
 FT TURN 486  
 FT STRAND 487  
 FT TURN 490  
 FT TURN 492  
 FT TURN 493  
 FT STRAND 496  
 FT STRAND 497  
 FT STRAND 501

Query Match 32.0%; Score 329.5; DB 1; Length 528;  
 Best Local Similarity 42.8%; Pred. No. 4.2e-23;  
 Matches 80; Conservative 28; Mismatches 66; Indels 13; Gaps 5;

QY 10 INFATAGATQSYNTNFTAVRGRLTTTCADVRHEIPVLPNVRVGLPINORFILLVLSNHAEL 69  
 DB 5 IKFTEGATQSYKQKFEALRRLRGG--LIHDPVLPDPTTLOERNRITVLSNSDTE 62  
 QY 70 SVTLALDVTNAYVVGVRAGNSAVFFH--PDNQEDAEATHLFTDVQNRITFAFGNYDRL 127

Db 63 SIEGVIDTWAYVAYRAGTOSYELRDAPSSASD-----YLFTGT-DQHSUPFGTYGDL 116  
QY 128 BOLAGNRELENGNPLEEBAISALYYSTGTQPTLARSFIIICIMISEARFOVIEG 187  
Db 117 ERWQHQRQIPLGLQALTHGIS---FFRSGNDNEEKARLIVIIQWVAARFRYSN 173  
QY 188 ERMTRIR 194  
Db 174 RRVRSIQ 180  
RESULT 10  
ID RIPI BRYDI STANDARD; PRT; 290 AA.  
AC P33185; Q9S819;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ribosome-inactivating protein bryodin I precursor (rRNA N-glycosidase)  
DE (EC 3.2.2.22) (BD1).  
OS Bryonia dioica (Red bryony).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.  
OX NCBI\_TaxID=3652;  
RN [1]  
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
RC TISSUE=Leaf;  
RX MEDLINE=97228081; PubMed=9115985;  
RA Gawlak S.L., Neubauer M., Klei H.E., Chang C.Y.Y., Einspahr H.M.,  
RA Siegall C.B.;  
RT "Molecular, biological, and preliminary structural analysis of  
RT recombinant bryodin I, a ribosome-inactivating protein from the plant  
RT Bryonia dioica.";  
RL Biochemistry 36:3095-3103(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Siegall C.B.;  
RT "Cloning and expression of a gene encoding bryodin I from Bryonia  
RT dioica.";  
RL Patent number US5541110, 30-JUL-1996.  
RN [3]  
RP SEQUENCE OF 24-66.  
RC TISSUE=Seed;  
RX MEDLINE=89326691; PubMed=2753596;  
RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Sorio M.,  
RA Lappi D.;  
RT "N-terminal sequence of some ribosome-inactivating proteins.";  
RL Int. J. Pept. Protein Res. 33:263-267(1989).  
RN [4]  
RP SEQUENCE OF 24-43.  
RC TISSUE=Root;  
RX MEDLINE=95151812; PubMed=7849072;  
RA Siegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,  
RA Marquardt H.;  
RT "Characterization of ribosome-inactivating proteins isolated from  
RT Bryonia dioica and their utility as carcinoma-reactive  
RT immunconjugates.";  
RL Bioconj. Chem. 5:423-429(1994).  
RN [5]  
RP FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS  
CC PROTEIN SYNTHESIS IN ANIMAL CELLS.  
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
CC specific adenosine on the 28S rRNA.  
CC -!- PTM: APPEARS TO UNDERGO PROTEOLYTIC CLEAVAGE IN THE C-TERMINAL TO  
CC PRODUCE A SHORTER PROTEIN.  
CC -!- BIOTECHNOLOGY: Especially useful as immunotoxin for  
CC pharmacological applications as it has low toxicity in rats and  
CC mice but is potent once inside target cells.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
CC TYPE 1 RIP SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; I24020; -, NOT\_ANNOTATED\_CDS.  
DR PIR; S16491; S16491.  
DR PDB; 1BRY; 04-MAR-98.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;  
KW 3D-structure; Multigene family; Glycoprotein; Signal.  
FT SIGNAL 1 23 RIBOSOME-INACTIVATING PROTEIN BRYODIN I.  
FT CHAIN 24 270 MISSING IN MATURE PROTEIN.  
FT PROPEP 271 290 BY SIMILARITY.  
FT ACT\_SITE 183 183  
FT ACT\_SITE 212 212  
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT MUTAGEN 212 212 E->K: REDUCES ACTIVITY 10-FOLD.  
FT CONFLICT 61 65 RSSIS -> LRHXI (IN REF. 3).  
FT STRAND 25 28  
FT TURN 30 31  
FT HELIX 34 46  
FT TURN 47 47  
FT STRAND 50 54  
FT TURN 55 56  
FT STRAND 57 60  
FT HELIX 66 69  
FT STRAND 70 76  
FT TURN 78 79  
FT STRAND 82 88  
FT TURN 89 92  
FT STRAND 93 99  
FT TURN 100 101  
FT STRAND 102 105  
FT HELIX 109 114  
FT TURN 115 117  
FT TURN 120 121  
FT STRAND 124 127  
FT HELIX 134 141  
FT TURN 142 142  
FT HELIX 145 147  
FT STRAND 150 150  
FT HELIX 152 163  
FT TURN 164 165  
FT HELIX 167 186  
FT STRAND 187 187  
FT HELIX 188 196  
FT STRAND 202 202  
FT HELIX 206 213  
FT TURN 214 214  
FT HELIX 215 225  
FT TURN 226 230  
FT STRAND 231 239  
FT TURN 241 242  
FT STRAND 245 250  
FT TURN 251 252  
FT HELIX 254 257  
FT TURN 258 259  
FT STRAND 260 260  
FT STRAND 263 263  
FT HELIX 266 268  
SQ SEQUENCE 290 AA; E965CD9C031A42DB CRC64;  
Query Match 30.3%; Score 312; DB 1; Length 290;  
Best Local Similarity 37.3%; Pred. No. 8.5e-22;  
Matches 69; Conservative 46; Mismatches 58; Indels 12; Gaps 5;

QY 10 INFRTAGATVQSYTNFIRAVRGLTTGADVREHPIVLPNVRVGLPINORFILVELSNHAE 69  
 DB 25 VSFRLSGATTTSYGVFIKNLREALPYERKV-YNIPLL--RSSISGSGRYTLHLHTNYADE 81  
 QY 70 SVTLALDVTNAVYVGVYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
 DB 82 TISVADVNTVNYMGVLGADVSVFF--NEASATEAKVFVDAKKKVTLPYSGNYERLQ 138  
 QY 129 QLAGNRENIELNGPLEEAISALYYSTGGTOLPTLARSFICIQMISAARFOVIEGE 188  
 DB 139 TAAGKIRENIPGLPALDSAITLYYYTAS-----SAASALLVLQSTABSAKYKIEQ 193  
 QY 189 MRTRI 193  
 DB 194 IGKRV 198

RESULT 11  
 RIBP\_LUCFY STANDARD; PRT; 250 AA.  
 AC P22851;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein luffin-B (rRNA N-glycosidase)  
 DE (EC 3.2.2.22).  
 OS Luffa cylindrica (Smooth loofah) (Sponge gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Luffa.  
 OX NCBI\_TaxID=3670;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=91248488; PubMed=1368666;  
 RA Islam M.R., Hirayama H., Funatsu G.;  
 RT "Complete amino acid sequence of luffin-b, a ribosome-inactivating  
 protein from sponge gourd (Luffa cylindrica) seeds.";  
 RL Agric. Biol. Chem. 55:229-238 (1991).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.

QY 10 INFRTAGATVQSYTNFIRAVRGLTTGADVREHPIVLPNVRVGLPINORFILVELSNHAE 69  
 DB 3 VSFSLGSKSYKSFITAKLALPSKEKVSNIPLLPASGA---SRVILMQLSNYDAK 59  
 QY 70 SVTLALDVTNAVYVGVYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLEQ 129  
 DB 60 AITMAIDVTNVYMGVLGADVSVFF--ANESDAKLASQYVFGKSTLVTPYSGNYERLQ 116  
 QY 130 LAGNRENIELNGPLEEAISALYYSTGGTOLPTLARSFICIQMISAARFOVIEGEM 189  
 DB 117 AAGKIREKIPGLFALDSALTSIFHYDS-----TAAAFVLVLTQTAESRKYKIEQ 171  
 QY 190 FTRIRYN 196  
 DB 172 IERIPKN 178

Query Match 29.2%; Score 300.5; DB 1; Length 250;  
 Best Local Similarity 34.2%; Pred. No. 8.3e-21;  
 Matches 64; Conservative 48; Mismatches 64; Indels 11; Gaps 3;

QY 10 INFRTAGATVQSYTNFIRAVRGLTTGADVREHPIVLPNVRVGLPINORFILVELSNHAE 69  
 DB 28 VSFSLGSKSYKSFITAKLALPSKEKVSNIPLLPASGA---SRVILMQLSNYDAK 83  
 QY 69 LSVTLALDVTNAVYVGVYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
 DB 84 NTITMAVDVTNVYMGVLGADVSVFF--NETDAQLASKEVFGTKITLTPYSNGYQKLQ 140  
 QY 129 QLAGNRENIELNGPLEEAISALYYSTGGTOLPTLARSFICIQMISAARFOVIEGE 188  
 DB 141 SVAREKDSIPLGFMALDSALTSIYYDSRAPI-----AFLVLIQTAAEARVYKIEQ 195  
 QY 189 MRTRIRYN 197  
 DB 196 IIDRISVSK 204

RESULT 13

RESULT 12  
 RIBP\_CUCFI STANDARD; PRT; 286 AA.  
 AC Q9FRX4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative ribosome-inactivating protein precursor (rRNA N-glycosidase)  
 DE (EC 3.2.2.22).  
 OS Cucumis figareii.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.  
 OX NCBI\_TaxID=131071;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada T., Ohki S.T., Osaki T.;  
 RT "Cloning and analysis of a cDNA coding a putative ribosome-  
 RT inactivating protein from Cucumis figareii.";  
 RL Plant Biotechnol. 17:337-340 (2000).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.

QY 10 INFRTAGATVQSYTNFIRAVRGLTTGADVREHPIVLPNVRVGLPINORFILVELSNHAE 68  
 DB 28 VSFSLGSKSYKSFITAKLALPSKEKVSNIPLLPASGA---SRVILMQLSNYDAK 83  
 QY 69 LSVTLALDVTNAVYVGVYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
 DB 84 NTITMAVDVTNVYMGVLGADVSVFF--NETDAQLASKEVFGTKITLTPYSNGYQKLQ 140  
 QY 129 QLAGNRENIELNGPLEEAISALYYSTGGTOLPTLARSFICIQMISAARFOVIEGE 188  
 DB 141 SVAREKDSIPLGFMALDSALTSIYYDSRAPI-----AFLVLIQTAAEARVYKIEQ 195  
 QY 189 MRTRIRYN 197  
 DB 196 IIDRISVSK 204

Query Match 28.8%; Score 296.5; DB 1; Length 286;  
 Best Local Similarity 34.4%; Pred. No. 2.3e-20;  
 Matches 65; Conservative 47; Mismatches 64; Indels 13; Gaps 5;

QY 10 INFRTAGATVQSYTNFIRAVRGLTTGADVREHPIVLPNVRVGLPINORFILVELSNHAE 68  
 DB 28 VSFSLGSKSYKSFITAKLALPSKEKVSNIPLLPASGA---SRVILMQLSNYDAK 83  
 QY 69 LSVTLALDVTNAVYVGVYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
 DB 84 NTITMAVDVTNVYMGVLGADVSVFF--NETDAQLASKEVFGTKITLTPYSNGYQKLQ 140  
 QY 129 QLAGNRENIELNGPLEEAISALYYSTGGTOLPTLARSFICIQMISAARFOVIEGE 188  
 DB 141 SVAREKDSIPLGFMALDSALTSIYYDSRAPI-----AFLVLIQTAAEARVYKIEQ 195  
 QY 189 MRTRIRYN 197  
 DB 196 IIDRISVSK 204

RESULT 13

	Query Match	27.8%; Score 286; DB 1; Length 254;
	Best Local Similarity	38.4%; Pred. No. 1.9e-19;
	Matches	81; Conservative 36; Mismatches 50; Indels 38; Gaps 10.
QY	14 TAGATVQSYYTNFIRAVRGRLTGTGADVRHEIPVLNRRVGLPIN--QRFILVELSNHAEILSV	71
DB	9 THQTGGEEYFRFITLLRDVVSGS-FSNEPILL-RQSTIPVDQAORFLIVELTNGQQDSV	66
OY	72 TIALDVTWNAVYGVPAGNSAYVFHFPHDNODEAFITHTLTDVNRYTEAFGGNDYRLEOLA	131

```
QY 130 LAGNLENTELNGPLLEAISAALYYSTGTQTLPTLARSFIIQIMISEAARFQYIEGEM 189
DB 136 AGKIREKIPLGPPALDSAITTLFHYDS-----TAAAAAFLVLIQTAAEASRKYIEGQI 190
QY 190 RTRIRYNR 197
DB 191 IERISKNQ 198

RESULT 15
RIP2 MOMBA
ID RIP2_MOMBA STANDARD; PRT; 286 AA.
AC P29339;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ribosome-inactivating protein momordin II precursor (xRNA
DE N-glycosidase) (EC 3.2.2.22).
OS Momordica balsamina (Bitter melon) (Balsam pear).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3672;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=93027170; PubMed=1408771;
RA Ortigas M., Better M.;
RT "Momordin II, a ribosome inactivating protein from Momordica
RT balsamina, is homologous to other plant proteins.";
RL Nucleic Acids Res. 20:4662-4662(1992).
CC -|- CATALYTIC ACTIVITY: Endonucleolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -|- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z12175; CAA78166.1; -
CC PIR; S25560; S25560.
CC PDB; 1CF5; 07-JUN-99.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00161; RIP; 1.
CC PRINTS; PR00396; SHIGARICIN.
CC PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin; Signal;
KW 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 286 RIBOSOME-INACTIVATING PROTEIN MOMORDIN
FT ACT SITE 181 181 BY SIMILARITY.
FT SEQUENCE 286 AA; 32031 MW; 3B89FFIAB6B25986 CRC64;
SQ
Query Match 27.1%; Score 279; DB 1; Length 286;
Best Local Similarity 35.3%; Pred. No. 9.8e-19;
Matches 66; Conservative 42; Mismatches 67; Indels 12; Gaps 5;
QY 10 INFTTAGATVQSYTNFRAVRGLTGTGADVREHIEPVLNVRGLPINQRFILVELSNHAEL 69
DB 25 VNFEDLSTAKTYTKFIEDFRATLPFSHKV-YDIPLLYSTIS--DSRRFILDLTSYAYE 81
QY 70 SVTLALDVNAYVYGRAGNSAYFFHPDQEDABAEATHLFTDVQNYTFAFGNYDRLEQ 129
DB 82 TISVAIDVTNVYVAYTRDVSYFF---KESPEAYNILFKGTR-KITLPTGTGNYENLOT 137
QY 130 LAGNLENTELNGPLLEAISAALYYSTGTQTLPTLARSFIIQIMISEAARFQYIEGEM 189
```

```
DB 138 AAHKIRENIDLGLPALSSAITTLFYNA-----QSAPSALLVLIQTAAEASRKYIERHV 192
QY 190 RTRIRYN 196
DB 193 AKYVATN 199

Search completed: February 10, 2004, 16:23:26
Job time : 7.67161 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:00 ; Search time 26.6864 Seconds  
(without alignments)  
1933.961 Million cell updates/sec

Title: US-10-083-336A-10  
Perfect score: 1029  
Sequence: 1 MIFPKQYPIINFATTAGATVQ.....RFOYIEGEMTRIRYNRSA 200

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 23.3\*

1: sp\_archea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1021	99.2	541	10 Q41174	Q41174 ricinus com
2	401.5	39.0	580	10 Q94BW3	Q94BW3 cinnamomum
3	397.5	38.6	580	10 Q94BW4	Q94BW4 cinnamomum
4	397.5	38.6	581	10 Q94BW5	Q94BW5 cinnamomum
5	395.5	38.4	549	10 Q9FV22	Q9FV22 cinnamomum
6	350.5	34.1	563	10 Q04367	Q04367 sambucus ni
7	347.5	33.8	564	10 Q9AVR2	Q9AVR2 sambucus eb
8	344.5	33.5	528	10 Q06076	Q06076 abrus prec
9	340	33.0	289	10 Q94KE4	Q94KE4 trichosan
10	338	32.8	247	10 Q9LRE3	Q9LRE3 trichosan
11	338	32.8	289	10 Q41216	Q41216 trichosan
12	336.5	32.7	252	10 Q38760	Q38760 abrus prec
13	330.5	32.1	563	10 Q94582	Q94582 sambucus ni
14	330.5	32.1	563	10 Q8GT32	Q8GT32 sambucus ni
15	329.5	32.0	252	10 Q38761	Q38761 abrus prec
16	324	31.5	270	10 Q8LPV7	Q8LPV7 trichosan

17	323.5	31.4	251	10 Q96236	Q96236 abrus prec
18	322.5	31.3	251	10 Q96237	Q96237 abrus prec
19	319	31.0	565	10 Q04071	Q04071 sambucus ni
20	317	30.8	270	10 Q41611	Q41611 trichosan
21	316.5	30.8	251	10 Q96235	Q96235 abrus prec
22	314.5	30.6	547	10 Q9M6E9	Q9M6E9 abrus prec
23	310.5	30.2	566	10 Q04072	Q04072 sambucus ni
24	307.5	29.9	278	10 Q00980	Q00980 luffa cylin
25	291	28.3	570	10 Q41358	Q41358 sambucus ni
26	285	27.7	249	10 Q8LKQ5	Q8LKQ5 viscum albu
27	285	27.7	570	10 Q22415	Q22415 sambucus ni
28	283.5	27.6	604	10 Q9M654	Q9M654 polygonatu
29	283	27.5	251	10 Q8LKQ4	Q8LKQ4 viscum albu
30	282	27.4	254	10 Q8LKQ6	Q8LKQ6 viscum albu
31	280	27.2	264	10 Q9FSH2	Q9FSH2 momordica c
32	278	27.0	293	10 Q8S452	Q8S452 jatropa cu
33	277.5	27.0	565	10 Q8W243	Q8W243 viscum albu
34	277	26.9	286	10 Q9FUV7	Q9FUV7 momordica c
35	277	26.9	569	10 P93543	P93543 sambucus ni
36	275	26.7	531	10 Q8RXH6	Q8RXH6 viscum albu
37	273	26.5	249	10 Q8RXH7	Q8RXH7 viscum albu
38	272	26.4	286	10 Q41257	Q41257 momordica c
39	272	26.4	298	10 Q04358	Q04358 iris hollan
40	270	26.2	592	10 Q8W2E7	Q8W2E7 iris hollan
41	268	26.0	300	10 Q04356	Q04356 iris hollan
42	266.5	25.9	603	10 Q9M653	Q9M653 polygonatu
43	262.5	25.5	299	10 Q8GZN9	Q8GZN9 euphorbia s
44	261	25.4	293	10 Q8VYU0	Q8VYU0 jatropa cu
45	259	25.2	573	10 Q8W2E8	Q8W2E8 iris hollan

## ALIGNMENTS

## RESULT 1

Q41174	ID	Q41174	PRELIMINARY;	PRT;	541 AA.
AC	Q41174;	01-NOV-1996 (Tremblrel. 01, Created)			
DT	01-NOV-1996 (Tremblrel. 01, Last sequence update)				
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)				
DE	Protein A chain (EC 3.2.2.22) (rRNA N-glycosidase)				
DE	(Fragment).				
OS	Ricinus communis (Castor bean).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC	eurosids I; Malpighiales; Euphorbiaceae; Ricinus.				
OX	NCBI_TaxID=3988;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92338377; PubMed=1633311;				
RA	Roberts L.M., Tregear J.W., Lord J.M.;				
RT	"Molecular cloning of ricin."				
RL	Targeted Diagn. Ther. 7:81-97(1992).				
CC	-1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE				
CC	SPECIFIC ADENOSINE ON THE 28S RRNA.				
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.				
DR	EMBL; S40366; AB22582.1; -				
DR	HSSP; P02879; IBER6.				
DR	InterPro; IPR000772; Ricin_B_lectin.				
DR	InterPro; IPR001574; RIP.				
DR	InterPro; IPR001400; Somatotropin.				
DR	Pfam; PF00652; Ricin_B_lectin; 6.				
DR	Pfam; PF00161; RIP; 1.				
DR	PRINTS; PR00396; SHIGARICIN.				
DR	SMART; SM00458; RICIN; 2.				
DR	PROSITE; PS0231; Ricin B LECTIN; 2.				
DR	PROSITE; PS00275; SHIGA RICIN; 1.				
DR	PROSITE; PS00338; SOMATOTROPIN_2; 1.				
KW	Hydrolase; Toxin.				
FT	NON_TER				
SEQ	SEQUENCE	541 AA;	60281 MW;	257B2CDEF1F2E9D9	CRC64;

Query Match 99.2%; Score 1021; DB 10; Length 541;  
 Best Local Similarity 99.5%; Pred. No. 9.3e-88;  
 Matches 198; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 IPKQYPIINFTAGATVQSYNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILV 61  
 DB 1 IPKQYPIINFTAGATVQSYNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILV 60

QY 62 ELSNHAELSVTLALDVTNAYVGVYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 121  
 DB 61 ELSNHAELSVTLALDVTNAYVGVYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 120

QY 122 GNYDRLEQLAGNLRENIEGLNGPLBEAISALYYSTGGTQPLTLARSFFIICIMISEAR 181  
 DB 121 GNYDRLEQLAGNLRENIEGLNGPLBEAISALYYSTGGTQPLTLARSFFIICIMISEAR 180

QY 182 FOYIEGEMETRIYRNSA 200  
 DB 181 FOYIEGEMETRIYRNSA 199

RESULT 2  
 Q94BW3 ID Q94BW3 PRELIMINARY; PRT; 580 AA.  
 AC Q94BW3;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomin III precursor  
 DE (EC 3.2.2.22) (tRNA N-glycosidase).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang Q., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomin proteins and study of their expression  
 RT patterns";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AY039803; AAK82460.1;  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR Pfam; PF00161; Ricin\_B\_lectin; 6.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS0231; RICIN\_B\_LECTIN; 2.  
 DR Hydroxylase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 580  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIN III.  
 SQ SEQUENCE 580 AA; 64421 MW; 940D10F01E7FB558 CRC64;

Query Match 39.0%; Score 401.5; DB 10; Length 580;  
 Best Local Similarity 50.0%; Pred. No. 3e-29;  
 Matches 94; Conservative 28; Mismatches 59; Indels 7; Gaps 5;

QY 7 YPIINFTAGATVQSYNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILVSN- 65  
 DB 33 YQTVTFTTKNATKTSYQTFEALRAQLASGEE-PHGIPVWRDGSIVPDSKRFILVLSNW 91

QY 66 HAELSVTLALDVTNAYVGVYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFGNYD 125  
 DB 92 AADSPVTLAVDTNAYVAYVAYTGSQSFLENDPDP--PAENLLPDTK-RYTFPFGSGSYT 148

QY 126 RLEQLAGNLRENIEGLNGPLBEAISALYYSTGGTQPLTLARSFFIICIMISEARFYI 185  
 DB 149 DLERVAGERREILLGMDFPLENAISLWNL--NQQRALARSLLIVVIQWVAEAVRFRFI 206

QY 186 EGEWRTRI 193  
 DB 207 EYRVRESI 214

## RESULT 3

Q94BW4 ID Q94BW4 PRELIMINARY; PRT; 580 AA.  
 AC Q94BW4;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomin II precursor  
 DE (EC 3.2.2.22) (tRNA N-glycosidase).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang Q., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomin proteins and study of their expression  
 RT patterns";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AY039802; AAK82459.1;  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR PRINTS; PR00161; Ricin\_B\_lectin; 1.  
 DR SMART; SM00396; SHIGARICIN.  
 DR PROSITE; PS0231; RICIN\_B\_LECTIN; 2.  
 DR Hydroxylase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 580  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIN II.  
 SQ SEQUENCE 580 AA; 64265 MW; 37B4289ECCE0CBFF CRC64;

Query Match 38.6%; Score 397.5; DB 10; Length 580;  
 Best Local Similarity 49.5%; Pred. No. 7.1e-29;  
 Matches 93; Conservative 30; Mismatches 58; Indels 7; Gaps 5;

QY 7 YPIINFTAGATVQSYNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILVSN- 65  
 DB 33 YQTVTFTTKNATKTSYQTFEALRAQLASGEE-PHGIPVWRDGSIVPDSKRFILVLSNW 91

QY 66 HAELSVTLALDVTNAYVGVYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFGNYD 125  
 DB 92 AADSPVTLAVDTNAYVAYVAYTGSQSFLENDPDP--PAENLLPDTK-RYTFPFGSGSYT 148

QY 126 RLEQLAGNLRENIEGLNGPLBEAISALYYSTGGTQPLTLARSFFIICIMISEARFYI 185  
 DB 149 DLERVAGERREILLGMDFPLENAISLWNL--WTNINQQRALARSLLIVVIQWVAEAVRFRFI 206

RESULT 4  
 Q94BW5 ID Q94BW5 PRELIMINARY; PRT; 581 AA.  
 AC Q94BW5;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomin I precursor  
 DE (EC 3.2.2.22) (tRNA N-glycosidase).



OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang Q., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomin proteins and study of their expression  
 RT patterns";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: AY039801; AA82458.1; -  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 5.  
 DR PRINTS; PF00161; RIP; 1.  
 DR SMART; SM00458; SHIGARICIN.  
 DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
 KW Hydrolyase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 581  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIN I.  
 SQ SEQUENCE 581 AA; 64215 MW; 68F5FB8FBA3D196 CRC64;  
 Query Match 38.6%; Score 397.5; DB 10; Length 581;  
 Best Local Similarity 50.0%; Pred. No. 7.1e-29;  
 Matches 94; Conservative 27; Mismatches 60; Indels 7; Gaps 5;  
 QY 7 YPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRGLPINORFILVELSN- 65  
 Db 33 YQVTFITTKAKTSYQTFEALRAQLASGEE-PHGIPVWRERSTVPSKRFILVELSNW 91  
 QY 66 HAELSVTLALDVNTAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRVYTFAPGNGYD 125  
 Db 92 AADSPVTLAVDTNAYVAYVAYRTGSGFFLRDNPDP--PAIENLLPDTK-RYTFPFSGSYT 148  
 QY 126 RLQIAGNLRNIEINGPLERALSALYYSTGGTGLPTLARSFIICMISEARFOYI 185  
 Db 149 DLEGVAGERREIEILLGMDPLENAISALWISNL--NQORALARSLLVVIQMVAAVRPFI 206  
 QY 186 EGEWRTI 193  
 Db 207 EYVRGSI 214  
 RESULT 5  
 Q9FV22 PRELIMINARY; PRT; 549 AA.  
 AC Q9FV22  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Type II ribosome-inactivating protein cinnamomin (BC 3.2.2.22) (rRNA  
 DE N-glycosidase) (fragment).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Xie L., Liu W.-Y., Wang E.-D.;  
 RT "Molecular cloning of cinnamomin A-, B-chain and the expression,  
 RT purification, characterization and mutagenesis of the A-chain";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AF259548; AAF68978.2; -  
 DR HSSP; P02879; 2AAL.

DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 5.  
 DR PRINTS; PF00161; RIP; 1.  
 DR SMART; SM00458; SHIGARICIN.  
 DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
 KW Hydrolyase; Toxin.  
 FT NON\_TER 1  
 FT SEQUENCE 549 AA; 60648 MW; 02607FE607CA44B0 CRC64;  
 Query Match 38.4%; Score 395.5; DB 10; Length 549;  
 Best Local Similarity 50.0%; Pred. No. 1e-28;  
 Matches 94; Conservative 27; Mismatches 60; Indels 7; Gaps 5;  
 QY 7 YPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRGLPINORFILVELSN- 65  
 Db 1 YQVTFITTKAKTSYQTFEALRAQLASGEE-PHGIPVWRERSTVPSKRFILVELSNW 59  
 QY 66 HAELSVTLALDVNTAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRVYTFAPGNGYD 125  
 Db 60 AADSPVTLAVDTNAYVAYVAYRTGSGFFLRDNPDP--PAIENLLPDTK-RYTFPFSGSYT 116  
 QY 126 RLQIAGNLRNIEINGPLERALSALYYSTGGTGLPTLARSFIICMISEARFOYI 185  
 Db 117 DLEGVAGERREIEILLGMDPLENAISALWISNL--NQORALARSLLVVIQMVAAVRPFI 174  
 QY 186 EGEWRTI 193  
 Db 175 EYVRGSI 182  
 RESULT 6  
 O04367 PRELIMINARY; PRT; 563 AA.  
 AC O04367  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Ribosome inactivating protein precursor (EC 3.2.2.22) (rRNA  
 DE N-glycosidase).  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98112023; PubMed=9450339;  
 RA Van Damme E.J., Roy S., Barre A., Rouge P., Van Leuven F.,  
 RA Peumans W.J.;  
 RT "The major elderberry (Sambucus nigra) fruit protein is a lectin  
 RT derived from a truncated type 2 ribosome-inactivating protein";  
 RL Plant J. 12:1251-1260(1997).  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; U76524; AAC15886.1; -  
 DR HSSP; P02879; 2AAL.  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR PRINTS; PF00161; RIP; 1.  
 DR SMART; SM00458; SHIGARICIN.  
 DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGARICIN; 1.  
 KW Hydrolyase; Signal; Toxin.  
 FT SIGNAL 1 25  
 FT CHAIN 26 297  
 FT CHAIN 298 563  
 FT SEQUENCE 563 AA; 62336 MW; 3ED2B6C0BE796205 CRC64;  
 SQ

Qy	125	DRLEQAGNLRNRETELGNGLPEEALISALYYSTGGTQLP	TLARSFIIQMISEAAARFOY 18
Db	142	DNLETAACRTRESSELGNPLDGAITSLWY--DGG----	VARSLVLIQWPEAARFRY 194
Qy	185	IEGEMRTFIR 194	
Db	195	IEQEVRSLSQ 204	
RESULT 8			
Qy	Q06076	PRELIMINARY;	PRT; 528 AA.
Db	Q06076	PRELIMINARY;	PRT; 528 AA.
Qy	01-NOV-1996	(TREMBLrel. 01, Created)	
Db	01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
Qy	01-NAR-2003	(TREMBLrel. 23, Last annotation update)	
Db	01-NAR-2003	(TREMBLrel. 23, Last annotation update)	
Qy	Abrin-d (EC 3.2.2.22)	(rRNA N-glycosidase) (Fragment).	
Db	Abrin-d (EC 3.2.2.22)	(rRNA N-glycosidase) (Fragment).	
Qy	Abrus precatorius (Indian licorice)	(Crab's eye)	
Db	Abrus precatorius (Indian licorice)	(Crab's eye)	
Qy	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
Db	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
Qy	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
Db	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
Qy	eurosid I; Fabales; Fabaceae; Papilionoideae; Abrus.		
Db	eurosid I; Fabales; Fabaceae; Papilionoideae; Abrus.		
Qy	NCBI_TaxID=3816;		
Db	NCBI_TaxID=3816;		
Qy	SEQUENCE FROM N.A.		
Db	SEQUENCE FROM N.A.		
Qy	MEDLINE=93132798; PubMed=8421313;		
Db	MEDLINE=93132798; PubMed=8421313;		
Qy	Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;		
Db	Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;		
Qy	"Primary structure of three distinct isoforms determined by cDNA		
Db	"Primary structure of three distinct isoforms determined by cDNA		
Qy	sequencing: conservation and significance."		
Db	sequencing: conservation and significance."		
Qy	J. Mol. Biol. 229:263-267(1993).		
Db	J. Mol. Biol. 229:263-267(1993).		
Qy	-!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE		
Db	-!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE		
Qy	SPECIFIC ADENOSINE ON THE 28S RRNA.		
Db	SPECIFIC ADENOSINE ON THE 28S RRNA.		
Qy	-!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.		
Db	-!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.		
Qy	EMBL; M98346; AAA32626.1; -.		
Db	EMBL; M98346; AAA32626.1; -.		
Qy	HSP; P11140; 1ABR.		
Db	HSP; P11140; 1ABR.		
Qy	InterPro; IPR000772; Ricin_B_lectin.		
Db	InterPro; IPR000772; Ricin_B_lectin.		
Qy	InterPro; IPR001574; RIP.		
Db	InterPro; IPR001574; RIP.		
Qy	Pfam; PF00652; Ricin_B_lectin; 6.		
Db	Pfam; PF00652; Ricin_B_lectin; 6.		
Qy	Pfam; PF00161; RIP; 1.		
Db	Pfam; PF00161; RIP; 1.		
Qy	PRINTS; PR00396; SHIGARICIN.		
Db	PRINTS; PR00396; SHIGARICIN.		
Qy	SMART; SM00458; RICIN; 2.		
Db	SMART; SM00458; RICIN; 2.		
Qy	PROSITE; PS0231; RICIN_B_LECTIN; 2.		
Db	PROSITE; PS0231; RICIN_B_LECTIN; 2.		
Qy	PROSITE; PS00275; SHIGA_RICIN; 1.		
Db	PROSITE; PS00275; SHIGA_RICIN; 1.		
Qy	Hydrolase; Toxin.		
Db	Hydrolase; Toxin.		
Qy	NON_TER 1		
Db	NON_TER 1		
Qy	FT 528		
Db	FT 528		
Qy	SEQUENCE 528 AA; 58870 MW; 62ED42FB9FF60F8 CRC64;		
Db	SEQUENCE 528 AA; 58870 MW; 62ED42FB9FF60F8 CRC64;		
Qy	Query Match	33.5%; Score 344.5; DB 10; Length 528;	
Db	Query Match	33.5%; Score 344.5; DB 10; Length 528;	
Qy	Best Local Similarity	45.0%; Pred. No. 6.3e-24;	
Db	Best Local Similarity	45.0%; Pred. No. 6.3e-24;	
Qy	Matches	85; Conservative 24; Mismatches 71; Indels 9; Gaps 4;	
Db	Matches	85; Conservative 24; Mismatches 71; Indels 9; Gaps 4;	
Qy	6	QYPIINFTAGATVQSYTNFIRAVGRLLTGADVHEIPVLPRVGLPILNQRFILVELSN 65	
Db	1	QDOVIKFTTEGATSQSYKQFIALRQRLTGG--LHIDIPVLDPDPTVEERNRIITVELSN 58	
Qy	66	HAEISVTLALDVNAVVGVRAGNSAYFPHDPNQDAEAIHLFDVQNRVYTFAPGGNYD 120	
Db	59	SERESEVIGDIVNAVIVAYRAGSQSYFL--RDAPASIVYLFPGTQ-RVSLRFDGSYG 11	
Qy	126	RLLEQAGNLRNRETELGNGLPEEALISALYYSTGGTQLP	TLARSFIIQMISEAAARFOY 18
Db	115	DLERWAHQTRREISLGLQALTHAIS---FLRGSASNDDEKARTLTVIIQMASEAAARFRCI 17	
Qy	186	EGEMRTFIR 194	
Db	172	SNRVGSIR 180	
RESULT 9			
Qy	Q94KE4	PRELIMINARY;	PRT; 289 AA.
Db	Q94KE4	PRELIMINARY;	PRT; 289 AA.
Qy	AC	Q94KE4;	
Db	AC	Q94KE4;	
Qy	01-DEC-2001	(TREMBLrel. 19, Created)	
Db	01-DEC-2001	(TREMBLrel. 19, Created)	

01-DEC-2001 (TrEMBLrel. 19, last sequence update)  
 01-MAR-2003 (TrEMBLrel. 23, last annotation update)  
 Trichosanthin precursor (EC 3.2.2.22) (rRNA N-glycosidase).  
 TCS.  
 OS Trichosanthes kirilowii (Mongolian snake-gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 NCBI\_TaxID=3677;  
 [1]  
 SEQUENCE FROM N.A.  
 RA Yuan H., Wang L., Wang Y., An C., Chen Z.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOTHEROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: AF367252; AAK52960.1; -.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS; PRO0396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 DR Hydrolyase; Signal; Toxin.  
 KW SIGNAL  
 FT CHAIN 24 270  
 FT SEQUENCE 289 AA; 31706 MW; AGD5602549CA5657 CRC64;  
 RP  
 Query Match 33.0%; Score 340; DB 10; Length 289;  
 Best Local Similarity 39.5%; Pred. No. 7.2e-24;  
 Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;  
 QY 10 INFRTAGATVQSYNTFIRAVRGRLTTGADVREHPIVLPNVRVGLPINORFILVELSNHAEL 69  
 Db 25 VSRFLSGATSSSYGVFISNLRKALPNERKL-YDIPLL--RSLPGSQRYALVHLTNVADE 81  
 QY 70 SVTLALDVTNAYVGYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
 Db 82 TISVAIDVTNIVYMGYRAGDTSYFF--NEASATEAAKYVFKDKRKVTLPLYSNGYERLQ 138  
 QY 129 QLAGNLRNIELGNPLEEASALYYSTGGTQPLTLARSFIICMISEARFOYIEGE 188  
 Db 139 TAAGKIRENIPLGLPALDSAITTLFYNNAN-----SAASALMVLIQSTSEARFYIEQQ 193  
 QY 189 METRI 193  
 Db 194 IGRKV 198  
 RESULT 10  
 QY 09LRE3 PRELIMINARY; PRT; 247 AA.  
 AC Q9LRE3  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Trichobakin (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).  
 GN TBK.  
 OS Trichosanthes sp. Bac Kan 8-98.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 NCBI\_TaxID=118182;  
 [1]  
 SEQUENCE FROM N.A.  
 RA Phan Van Chi, Nguyen Thuy Ha, Le Tran Binh;  
 RT "Genomic DNA clone for mature typ-1 ribosome-inactivating protein from  
 RT Trichosanthes sp. sample 01 Bac Kan 8-98 Vien CINH (Hanoi).";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOTHEROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: AB039324; BAA92530.1; -.  
 DR HSSP; P09989; IMRJ.  
 DR InterPro; IPR001574; RIP.

DR Pfam: PF00161; RIP; 1.  
 DR PRINTS; PRO0396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolyase; Toxin.  
 FT NON TER 1  
 FT NON TER 247  
 SQ SEQUENCE 247 AA; 27199 MW; 89811AC32892F03F CRC64;  
 RP  
 Query Match 32.8%; Score 338; DB 10; Length 247;  
 Best Local Similarity 40.0%; Pred. No. 8.9e-24;  
 Matches 74; Conservative 46; Mismatches 53; Indels 12; Gaps 5;  
 QY 10 INFRTAGATVQSYNTFIRAVRGRLTTGADVREHPIVLPNVRVGLPINORFILVELSNHAEL 69  
 Db 2 VSRFLSGATSSSYGVFISNLRKALPNERKL-YDIPLL--RSLPGSQRYALVHLTNVADE 58  
 QY 70 SVTLALDVTNAYVGYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
 Db 59 TISVAIDVTNIVYMGYRAGDTSYFF--NEASATEAAKYVFKDKRKVTLPLYSNGYERLQ 115  
 QY 129 QLAGNLRNIELGNPLEEASALYYSTGGTQPLTLARSFIICMISEARFOYIEGE 188  
 Db 116 IARAGKIRENIPLGLPALDSAITTLFYNNAN-----SAASALMVLIQSTSEARFYIEQQ 170  
 QY 189 METRI 193  
 Db 171 IGRKV 175  
 RESULT 11  
 QY 041216 PRELIMINARY; PRT; 289 AA.  
 AC Q41216  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Trichosanthin (EC 3.2.2.22) (rRNA N-glycosidase).  
 GN TRICHOSANTHIN, TCS.  
 OS Trichosanthes kirilowii (Mongolian snake-gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 NCBI\_TaxID=3677;  
 [1]  
 SEQUENCE FROM N.A.  
 RA MEDLINE=94271613; PubMed=8003348;  
 RA Zheng H., Wang B., Shaw P., Yeung H.;  
 RT "[Cloning and DNA sequencing of the gene encoding trichosanthin].";  
 RL I Chuan Heueh Pac 21:42-51(1994).  
 CC -!- CATALYTIC ACTIVITY: ENDOTHEROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: S70176; AB31048.1; -.  
 DR HSSP; P09989; IMRJ.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS; PRO0396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolyase; Toxin.  
 SQ SEQUENCE 289 AA; 31650 MW; 286AC14D48BCA175 CRC64;  
 RP  
 Query Match 32.8%; Score 338; DB 10; Length 289;  
 Best Local Similarity 39.5%; Pred. No. 1.1e-23;  
 Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;  
 QY 10 INFRTAGATVQSYNTFIRAVRGRLTTGADVREHPIVLPNVRVGLPINORFILVELSNHAEL 69  
 Db 25 VSRFLSGATSSSYGVFISNLRKALPNERKL-YDIPLL--RSLPGSQRYALVHLTNVADE 81  
 QY 70 SVTLALDVTNAYVGYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
 Db 82 TISVAIDVTNIVYMGYRAGDTSYFF--NEASATEAAKYVFKDKRKVTLPLYSNGYERLQ 138

QY 129 QIAGNLRNIELNGPLBEAISALYYSTGGTQTLPLARSFFICQMISEARFOYIEGE 188  
 Db 139 TAAGKIRENIPGLPALDSAITTFYNNAN-----SAASALMVLIQSTSEARFYKFIQQ 193  
 QY 189 MKTRI 193  
 Db 194 IGRV 198

RESULT 12  
 Q38760 PRELIMINARY; PRT; 252 AA.  
 AC Q38760;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE ABRIN-E (RRNA N-glycosidase) (EC 3.2.2.22) (Fragment).  
 GN RIP.  
 OS Abrus precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 NCBI\_TaxID=3816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEAF;  
 RX MEDLINE=91201329; PubMed=2016300;  
 RA Evensen G., Mathiesen A., Sundan A.;  
 RT "Direct molecular cloning and expression of two distinct abrin A-  
 chains";  
 RL J. Biol. Chem. 266:6848-6852 (1991).  
 CC -!- FUNCTION. THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.  
 CC ABRIN-A IS MORE TOXIC THAN RICIN.  
 CC -!- FUNCTION. THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
 FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
 PRECEDES ENDOCYTOSIS.  
 CC -!- CATALYTIC ACTIVITY. ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SUBUNIT. DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -!- DOMAIN. THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -!- SIMILARITY. TO OTHER BACTERIAL AND PLANTS RIBOSOME-INACTIVATING  
 PROTEINS. BELONGS TO TYPE 2 RIP.  
 DR EMBL; X54872; CAA38654.1; -;  
 DR HSP; P11140; IABR.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydroxylase; Glycosidase; Toxin; Repeat; Glycoprotein; Lectin.  
 FT CHAIN 1 252 ABRIN E, A CHAIN (BY SIMILARITY).  
 FT NON\_TER 252  
 SQ SEQUENCE 252 AA; 28309 MW; BBFC846B9E92B5DE CRC64;

Query Match 32.7%; Score 336.5; DB 10; Length 252;  
 Best Local Similarity 44.9%; Pred. No. 1.3e-23;  
 Matches 83; Conservative 24; Mismatches 69; Indels 9; Gaps 4;

QY 10 INFETAGATQSYTNFIRAVRGRLTTGADVRHEIPVLPNVRGLPINQRFILVELSNHAE 69  
 Db 6 IKFSTEGATSQYQFIEALRERLGG--LIHDPVLRDPTTVEERNYITVELSNRE 63  
 QY 70 SVTLALDVTNAYVGYRAGNSAYFFHPDQEDAEAI--THLFTDVQNRYYTFAFGNYYDLEQ 129  
 Db 64 SIEVGIDVTNAYWAYRAGSOSYFL---RDAPASASTYLTGTQ--RYSLRFDGSGYGLER 119  
 QY 130 LAGNLRNIELNGPLBEAISALYYSTGGTQTLPLARSFFICQMISEARFOYIEGEM 189  
 Db 120 WAHQTRQLSUGLQALHTAIS---FLRSGANDEKARTILVLIOMASEAARYIYNRV 176  
 QY 190 RTRIR 194

Db 177 GVSIR 181

RESULT 13  
 Q94592 PRELIMINARY; PRT; 563 AA.  
 AC Q94592;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Ribosome-inactivating protein (EC 3.2.2.22) (rRNA  
 N-glycosidase).  
 GN AVL  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RA Van Damme E.J.M.;  
 RT "Characterization and cloning of lectins and ribosome-inactivating  
 proteins from Sambucus nigra leaves";  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY. ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SIMILARITY. BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AF409135; AAL04123.1; -;  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydroxylase; Toxin.  
 SQ SEQUENCE 563 AA; 62242 MW; 07F7CBEDCF33BF10 CRC64;

Query Match 32.1%; Score 330.5; DB 10; Length 563;  
 Best Local Similarity 39.5%; Pred. No. 1.4e-22;  
 Matches 75; Conservative 39; Mismatches 61; Indels 15; Gaps 4;

QY 7 YPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRHEIPVLPNVRGLPINQRFILVELSNH 66  
 Db 28 YPSVSNLDGAKSATYRDFLSNLRKTVATGTYEVNGLPVLRESEYQVKSFRVFLVLTNY 87  
 QY 67 AELSVTLALDVTNAYVGYRAGNSAYFFHPDQEDAEAI--THLFTDVQNRYYTFAFGNYY 124  
 Db 88 NGMTVTILAVDVTNLYVVAESGNANSYFF-----KDATEVQKSNLFVGTQKN--TISFTGNY 141  
 QY 125 DRLEQLAGNLRNIELNGPLBEAISALYYSTGGTQTLPLARSFFICQMISEARFOY 184  
 Db 142 DNLTEAANTRESIELGPGPLDGAITSYLVHGD-----SVARSLLVLIQWSEARFRY 194  
 QY 185 IEQEMRTRIR 194  
 Db 195 IEQEVRSIQ 204

RESULT 14  
 Q8GT32 PRELIMINARY; PRT; 563 AA.  
 AC Q8GT32;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein nigrin 1 precursor  
 (EC 3.2.2.22).  
 GN Sambucus nigra (European elder).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;



1	955	97.9	267	14	AAR37290	Un
2	955	97.9	267	16	AAR63902	Ricin A-chain (RTA)
3	955	97.9	290	18	AAW25136	Ricin A-chain ribo
4	955	97.9	290	18	AAW21699	Ricin A-chain R1P.
5	955	97.9	332	8	AAP70097	Ricin A. Escheric
6	955	97.9	332	8	AAP70838	Sequence of Ricinu
7	955	97.9	332	10	AAP95639	Ricin A encoded by
8	955	97.9	554	16	AAR70827	Anti-catacar immu
9	955	97.9	562	10	AAP90079	Ricin D. Ricinu

XX Type II ribosome-inactivating protein; type II RIP; gelonin;  
 KW reomardin; immunosuppressant; autoimmune disease; cell killing; toxin;  
 YW reomardin; immunosuppressant; autoimmune disease; cell killing; toxin;

PT Analogues of type I ribosome inactivating protein - useful as  
 PT cytotoxic agents, immuno toxins for treating autoimmune diseases,  
 PT cancer, graft versus host disease and selective cell killing in-vivo  
 XX  
 PS Claim 1; Page 92; 163pp; English.

XX The invention covers analogues of Type I RIPs. Ricin is a Type II  
 CC RIP whose A chain is homologous to plant type I RIPs. The analogues  
 CC of the invention have a cysteine available for intermolecular  
 CC disulphide bonding at an amino acid position corresp. to a position  
 CC not naturally available for bonding; the cysteine residue is located in  
 CC the C-terminal region of the analogue between a position corresp. to  
 CC amino acid 251 and the C-terminus of ricin A chain. The analogues are  
 CC pref. joined via a disulphide linkage to a molecule which specifically  
 CC binds to a target cell, e.g. an antibody fragment.  
 CC (Updated on 09-JAN-2003 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 267 AA;

Query Match 97.9%; Score 955; DB 14; Length 267;  
 Best Local Similarity 95.0%; Pred. No. 3.7e-94;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFATTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51  
 DB 1 IFPKQYPIINFATTAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 60  
 QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFPHDPNQEDAEAIHTLFTDVQNRVTFAG 111  
 DB 61 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFPHDPNQEDAEAIHTLFTDVQNRVTFAG 120  
 QY 112 GNYDRLEQLAGNLRENIEELNGPLLEAISALYYVYSTGGTQPLTARSFFIICQMISEAAR 171  
 DB 121 GNYDRLEQLAGNLRENIEELNGPLLEAISALYYVYSTGGTQPLTARSFFIICQMISEAAR 180  
 QY 172 FOYIEGEMTRIRYNRNSA 190  
 DB 181 FOYIEGEMTRIRYNRNSA 199

RESULT 2  
 AAR63902  
 ID AAR63902 standard; protein; 267 AA.

XX AAR63902;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 27-JUL-1995 (first entry)  
 XX Ricin A-chain (RTA).  
 DE  
 XX Ricin A chain; RTA; ribosome-inactivating proteins; RIPs;  
 KW cytotoxic therapeutic agents; autoimmune disease; cancer;  
 KW graft-versus-host disease.  
 XX Ricinus communis.  
 XX OS  
 XX W09426910-A1.  
 XX  
 XX 24-NOV-1994.  
 XX  
 XX 12-MAY-1994; 94WO-US05348.  
 XX  
 XX 12-MAY-1993; 93US-0064691.  
 XX (XOMA) XOMA CORP.  
 XX  
 XX Better MD, Carroll SS, Studnicka GM, Carroll SF;  
 PI WPI; 1995-006804/01.  
 XX  
 XX Polynucleotide(s) encoding type I ribosome-inactivating proteins

PT - which are suitable for use as components of cytotoxic  
 PT therapeutic agents.

XX Example 3; Fig 1; 221pp; English.

XX AAR63902 is the ricin A chain gene product, it is analogous to the  
 CC ribosome-inactivating proteins (RIPs) described in AAR63903-R63911.  
 CC RIPs are the key components of cytotoxic therapeutic agents (CTAs),  
 CC which include gene fusion products and immunoconjugates. CTAs may  
 CC be used to selectively eliminate any cell type to which a RIP  
 CC component is targeted, by the specific binding capacity of the  
 CC second component of the agent. They can be used in the treatment  
 CC of diseases where the elimination of a particular cell type is  
 CC desired, such as autoimmune disease, cancer and graft-versus-host  
 CC disease.  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 267 AA;

Query Match 97.9%; Score 955; DB 16; Length 267;  
 Best Local Similarity 95.0%; Pred. No. 3.7e-94;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFATTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51  
 DB 1 IFPKQYPIINFATTAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 60  
 QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFPHDPNQEDAEAIHTLFTDVQNRVTFAG 111  
 DB 61 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFPHDPNQEDAEAIHTLFTDVQNRVTFAG 120  
 QY 112 GNYDRLEQLAGNLRENIEELNGPLLEAISALYYVYSTGGTQPLTARSFFIICQMISEAAR 171  
 DB 121 GNYDRLEQLAGNLRENIEELNGPLLEAISALYYVYSTGGTQPLTARSFFIICQMISEAAR 180  
 QY 172 FOYIEGEMTRIRYNRNSA 190  
 DB 181 FOYIEGEMTRIRYNRNSA 199

RESULT 3  
 AAW25136  
 ID AAW25136 standard; Protein; 290 AA.

XX AAW25136;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 02-DEC-1997 (first entry)  
 XX Ricin A-chain ribosome inhibitory protein inactive precursor.  
 DE  
 XX Maize; proRIP; ribosome inactivating protein; alpha; beta subunit;  
 KW internal linker; Barley Translation Inhibitor; Trichosanthin;  
 KW Ricin A-chain; Abrin-A A-chain; Saporin; SLT-1; Luffin A; MAP;  
 KW Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;  
 KW therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;  
 KW post-translational modification; cancer; neoplasia; HIV; AIDS;  
 KW human immunodeficiency virus; acquired immune deficiency syndrome.  
 XX  
 XX Synthetic.  
 XX OS  
 XX US5646026-A.  
 XX  
 XX 08-JUL-1997.  
 XX  
 XX 07-JUN-1995; 95US-0485286.  
 XX  
 XX 09-DEC-1992; 92US-0987927.  
 XX 11-JUN-1990; 90US-0535636.  
 XX 26-JAN-1995; 95US-0378761.  
 XX 07-JUN-1995; 95US-0485286.  
 XX (DOWC) DOWELANCO.

XX Hey TD, Morgan AER, Walsh TA;  
 XX WPI; 1997-362934/33.  
 XX  
 XX DNA encoding pro-ribosome inactivating proteins - inactive  
 FT precursors of ribosome inactivating proteins; can be expressed in  
 PT eukaryotic cells without causing cell death  
 XX  
 XX Claim 4; Column 91-94; 186pp; English.

XX AAW25136 represents a Ricin A-chain ribosome inhibitory protein (RIP)  
 CC which was engineered to contain a selectively removable internal peptide  
 CC linker sequence separating the alpha and beta units of the RIP. When  
 CC separated the two units regain activity and are capable of inactivating  
 CC eukaryotic ribosomes and hence preventing protein production. Many  
 CC different RIPs may be produced with an internal linker including  
 CC maize RIP, Trichosanin, Ricin A-chain, Abrin-A A-chain and  
 CC Saporin. The RIPs can be used in the construction of therapeutic  
 CC toxins targeted to specific cells such as tumour cells via the  
 CC attachment of a targeting polypeptide, e.g. a monoclonal antibody.  
 CC A further use is in HIV therapy (see US4869903). There is interest  
 CC in expressing RIP recombinantly in host eukaryotic cells, because of  
 CC the capacity to provide correct post-translational processing. However,  
 CC RIPs effectively inhibit protein synthesis in eukaryotic cells resulting  
 CC in cell death. Since the inactive RIP proteins are not cytotoxic to  
 CC eukaryotic cells, they can be recombinantly expressed in such cells and  
 CC then converted to active RIP proteins.  
 CC (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 290 AA;

Query Match 97.9%; Score 955; DB 18; Length 290;  
 Best Local Similarity 95.0%; Pred. No. 4.le-94;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLT-----VLPNRVGLPINQRFILV 51  
 DB 25 IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTGDVVRHEIPVLPNRVGLPINQRFILV 84  
 QY 52 ELSNHAELSVTLDVNTAVVGVGRAGNSAYFFHPDQEDAEATHLFTDQVQRYTFAFG 111  
 DB 85 ELSNHAELSVTLDVNTAVVGVGRAGNSAYFFHPDQEDAEATHLFTDQVQRYTFAFG 144  
 QY 112 GNYDRLEQLAGNRENELNGNPLEEASISALYYSTGCTQPLTARSFIICIMISEAAR 171  
 DB 145 GNYDRLEQLAGNRENELNGNPLEEASISALYYSTGCTQPLTARSFIICIMISEAAR 204  
 QY 172 FOYIEGEMRTRIYNRNSA 190  
 DB 205 FOYIEGEMRTRIYNRNSA 223

RESULT 4

AAW21699

ID AAW21699 standard; Protein; 290 AA.

XX AC

XX AAW21699;

XX 25-MAR-2003 (updated)

DT 26-SEP-1997 (first entry)

XX Ricin A-chain RIP.

XX pro-Ribosome inactivating protein; proRIP; peptide linker; cancer;  
 KW inactivation; eukaryotic ribosome; alpha fragment; beta fragment;  
 KW inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver;  
 KW rat; ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.

OS Ricinus communis.

XX Key Location/Qualifiers

XX Region 152...162

FT

/note= "Position of possible insertion of internal  
 peptide linker sequence"

US5635384-A.

03-JUN-1997.

26-JAN-1995; 95US-0378761.

09-DEC-1992; 92US-0987927.

11-JUN-1990; 90US-0535636.

26-JAN-1995; 95US-0378761.

(DOMC ) DOWELANCO.

Hey TD, Morgan AER, Walsh TA;

WPI; 1997-309831/28.

Inactive precursor of maize ribosome-inactivating protein - also

chimeric ribosome-inactivating protein precursors containing

internal linker sequences

Claim 2; Column 91-94; 121pp; English.

The sequences given in AAW21698-710 represent Ribosome Inactivating  
 Proteins (RIP's), which may be used in the construction of the  
 proRIP of the invention. The proRIP has a selectively removable,  
 internal peptide linker. The precursor sequence is incapable of  
 inactivating eukaryotic ribosomes, but can be converted by removal  
 of the linker into a protein having alpha and beta fragments and being  
 capable of inactivating eukaryotic ribosomes. RIPs are potent  
 inhibitors of eukaryotic protein synthesis. They possess a highly  
 specific N-glycosidase activity which cleaves the glycosidic bond of  
 adenine 4324 of rat liver ribosomal 28S RNA. RIP's selectively inhibit  
 cellular proliferation of cells, e.g. cancer cells and HIV-infected T  
 cells. The inactive proRIP proteins make it possible to provide protein  
 synthesis inhibitors with uses in practical and improved ways not before  
 possible. The RIP can be used to make cytotoxic conjugates.  
 (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 290 AA;

Query Match 97.9%; Score 955; DB 18; Length 290;

Best Local Similarity 95.0%; Pred. No. 4.le-94;

Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLT-----VLPNRVGLPINQRFILV 51

DB 25 IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTGDVVRHEIPVLPNRVGLPINQRFILV 84

QY 52 ELSNHAELSVTLDVNTAVVGVGRAGNSAYFFHPDQEDAEATHLFTDQVQRYTFAFG 111

DB 85 ELSNHAELSVTLDVNTAVVGVGRAGNSAYFFHPDQEDAEATHLFTDQVQRYTFAFG 144

QY 112 GNYDRLEQLAGNRENELNGNPLEEASISALYYSTGCTQPLTARSFIICIMISEAAR 171

DB 145 GNYDRLEQLAGNRENELNGNPLEEASISALYYSTGCTQPLTARSFIICIMISEAAR 204

QY 172 FOYIEGEMRTRIYNRNSA 190

DB 205 FOYIEGEMRTRIYNRNSA 223

RESULT 5

AAW70097

ID AAW70097 standard; protein; 332 AA.

XX AC

XX AAW70097;

XX 09-APR-1991 (first entry)

XX DT

XX Ricin A.



```

XX KW Ricin A; Met-aminopeptidase.
XX OS Escherichia coli.
XX PN EP219237-A.
XX PD 22-APR-1987.
XX PF 19-SEP-1986; 86EP-0307242.
XX PR 06-MAY-1986; 86US-0860330.
XX PR 20-SEP-1985; 85US-0778414.
XX PR (CETU ) CETUS CORP.
XX PA Benbassat A, Bauer KA, Chang S, Chang SY;
XX PI WPI; 1987-110172/16.
XX DR N-PSDB; AAN70152.
XX CC N-terminal methionine free proteins prodn. - by using host
XX CC transformed with vector to express a methionine-amino-peptidase
XX PT Disclosure; Fig. 4; 20pp; English.
XX PS Ricin A may be produced in a form which lacks an N-terminal Met
XX CC using Met-aminopeptidase from E.coli.
XX CC
XX SQ Sequence 332 AA;
XX Query Match 97.9%; Score 955; DB 8; Length 332;
XX Best Local Similarity 95.0%; Pred. No. 4.9e-94;
XX Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
QY 2 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRIT-----VLPNRVGLPINQRFILV 51
DB 36 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRITGADVVRHEIPVLPNRVGLPINQRFILV 95
QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNQEDAEAIHTLFTDVQNRVTFAPG 111
DB 96 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNQEDAEAIHTLFTDVQNRVTFAPG 155
QY 112 GNYDRLEQLAGNLRENIELNGPLERISALYYSTGGTQLPTLARSFICIQMISEAAR 171
DB 156 GNYDRLEQLAGNLRENIELNGPLERISALYYSTGGTQLPTLARSFICIQMISEAAR 215
QY 172 FOYIEGEMTRIRYNRRA 190
DB 216 FOYIEGEMTRIRYNRRA 234
RESULT 6
AAP70838
ID AAP70838 standard; protein; 332 AA.
XX AAP70838;
XX 25-MAR-2003 (updated)
XX 18-FEB-1991 (first entry)
XX Sequence of Ricinus communis castor beans ricin toxin (RT or ricin)
XX A protein encoded by pR123.
XX Lectin; toxin protein; cytotoxic; cytostatic; castor bean;
XX plant toxin.
XX Ricinus communis.
XX Key Location/Qualifiers
XX Region 1..32
XX FT /note="Leader"
XX FT 33..302
XX FT

```

---

```

FT FT /note="A-chain"
FT FT 315..332
FT FT /note="B-chain"
XX EP237676-A.
XX 23-SEP-1987.
XX 13-NOV-1986; 86EP-0308877.
XX 07-MAR-1986; 86US-0837583.
XX (CETU ) CETUS CORP.
XX (CHIR ) CHIRON CORP.
XX PI Piatak M;
XX WPI; 1987-265177/38.
XX DR N-PSDB; AAN70519.
XX CC The full-length sequences encoding ricin A (AAN70520), ricin D
XX CC (AAN70525), putative ricin E (AAN70526) and RCA (AAN70524) in precursor
XX CC then were obtd. using messenger RNA to obtain a cDNA library, and
XX CC then probing the library to retrieve the desired cDNA inserts. The
XX CC library was probed using the 35-mer given in AAN70514. Figure 4 (see
XX CC AAN70520, AAN70521, AAN70522) shows the nucleotide sequences of three
XX CC plasmids contg. cDNA inserts obtd. by probing a cDNA library for
XX CC sequences encoding ricin B using the probe in AAN70517. The cDNA
XX CC inserts can be placed into expression vectors. Site-directed
XX CC mutagenesis may be used to place an ATG start codon and a HindIII
XX CC site at the beginning of the mature protein (see AAN70518). The
XX CC coding sequences of the inserts can be ligated into expression
XX CC vectors contg. the PhOA promoter-operator and leader sequence
XX CC (AAN70523) and suitable retroregulators.
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 332 AA;
XX Query Match 97.9%; Score 955; DB 8; Length 332;
XX Best Local Similarity 95.0%; Pred. No. 4.9e-94;
XX Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
QY 2 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRIT-----VLPNRVGLPINQRFILV 51
DB 36 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRITGADVVRHEIPVLPNRVGLPINQRFILV 95
QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNQEDAEAIHTLFTDVQNRVTFAPG 111
DB 96 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNQEDAEAIHTLFTDVQNRVTFAPG 155
QY 112 GNYDRLEQLAGNLRENIELNGPLERISALYYSTGGTQLPTLARSFICIQMISEAAR 171
DB 156 GNYDRLEQLAGNLRENIELNGPLERISALYYSTGGTQLPTLARSFICIQMISEAAR 215
QY 172 FOYIEGEMTRIRYNRRA 190
DB 216 FOYIEGEMTRIRYNRRA 234
RESULT 7
AAP95639
ID AAP95639 standard; protein; 332 AA.
XX AAP95639;
XX AC AAP95639;
XX 25-MAR-2003 (updated)
XX 31-OCT-2002 (updated)
XX DT

```

DT 13-AUG-1990 (first entry)  
XX Ricin A encoded by insert from plasmid pRA123.  
XX Plasmid pRA123; ricin-A; ricin-B; cytotoxicity.  
XX Ricinus communis.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH Peptide 1..35  
FT /label= leader sequence  
FT Peptide 36..302  
FT /label=A-chain  
FT Peptide 303..314  
FT /label=linker  
FT Peptide 315..332  
FT /label=B-chain  
XX EP335476-A.  
PN  
XX  
XX 04-OCT-1989.  
XX  
XX 19-JAN-1989; 89EP-0201162.  
XX  
XX 08-FEB-1984; 84US-0578115.  
PR 08-FEB-1984; 84US-0578121.  
PR 09-FEB-1984; 84US-0578122.  
PR 07-SEP-1984; 84US-0648759.  
PR 20-SEP-1984; 84US-0653515.  
XX  
XX (CETU ) CETUS CORPORATION.  
XX  
XX Gelfand D, Lawyer FC, Horn G, Greenfield L, Nitecki D, Kaplan D;  
PI Piatak MJ;  
XX  
XX WPI; 1989-286959/40.  
DR N-PSDB; AAN91281.  
XX  
XX Recombinant vectors expressing ricin chains or diphtheria toxin -used for  
FT prodn. of new immunotoxin conjugates with monoclonal antibodies, having  
FT high cell specificity and good extracellular stability.  
XX  
XX Disclosure; Fig 14; 54pp; English.  
XX  
XX Plasmid RA123 (ATCC No. 39799) carries the entire coding sequence for  
CC ricin A, as well as codons for 12 AAs joining the A to the B chain.  
CC Following modification for ease of manipulation the plasmid was used to  
CC construct expression vectors which express the conjugates in  
CC host cells.  
CC (Updated on 31-OCT-2002 to add missing OS field.)  
CC (Updated on 25-MAR-2003 to correct PF field.)  
CC (Updated on 25-MAR-2003 to correct PR field.)  
CC (Updated on 25-MAR-2003 to correct PI field.)  
XX  
SQ Sequence 332 AA;  
Query Match 97.9%; Score 955; DB 10; Length 332;  
Best Local Similarity 95.0%; Pred. No. 4.9e-94;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINTTAGATVQSYTNFRAVRGLT-----VLPNRVGLPINORFLV 51  
DB 36 IFPKQYPIINTTAGATVQSYTNFRAVRGLTGCADVHRHPIVLPNVRVGLPINORFLV 95  
QY 52 ELSNHAELSVTIALDVTNAYVGVYAGNSAYFFHPDNCQEDAEATHLFTDVQNYTFAFG 111  
DB 96 ELSNHAELSVTIALDVTNAYVGVYAGNSAYFFHPDNCQEDAEATHLFTDVQNYTFAFG 155  
QY 112 GNYDRLEQAGLNRENIELGNGLPEEALISALYYSTGTQTLPTARSFIICQIMISEAAR 171  
DB 156 GNYDRLEQAGLNRENIELGNGLPEEALISALYYSTGTQTLPTARSFIICQIMISEAAR 215

QY 172 FOYIEGEMTRIRYRRSA 190  
DB 216 FOYIEGEMTRIRYRRSA 234  
RESULT 8  
AAR70827  
ID AAR70827 standard; Protein; 554 AA.  
XX  
XX AAR70827;  
XX AC  
XX 25-MAR-2003 (updated)  
DT 31-AUG-1995 (first entry)  
XX  
XX Anti-cataract immunotoxin.  
XX Immunotoxin; heavy chain; light chain; variable region; antibody;  
KW ricin-A; cytostatic; cataract; lens opacification; epithelial cell;  
KW pHBL19; 4197X; monoclonal antibody; MAB.  
XX Synthetic.  
OS  
XX Location/Qualifiers  
FH Key 1..27  
FT Peptide /label= Sig\_peptide  
FT /note= "phoA signal sequence"  
FT Domain 28..145  
FT /label= HEAVY  
FT /note= "Mab 4197X heavy chain"  
FT Peptide 148..166  
FT /label= LINKER  
FT Domain 169..274  
FT /label= LIGHT  
FT /note= "Mab 419X light chain"  
FT Domain 276..544  
FT /label= RICIN-A  
FT Peptide 549..554  
FT /label= TAG  
FT /note= "hexa-histidine tail"  
XX  
XX WO9503828-A1.  
PN  
XX  
XX 09-FEB-1995.  
XX  
XX 15-JUL-1994; 94WO-US07919.  
XX  
XX 02-AUG-1993; 93US-0101329.  
XX  
XX (HOUS-) HOUSTON BIOTECHNOLOGY INC.  
XX  
XX Gould RM, Kelleher PJ, Wallace TL, Wood MS;  
XX WPI; 1995-082036/11.  
DR N-PSDB; AAQ85386.  
XX  
XX New single chain immuno-toxin - binds specifically to epithelial  
FT cells, for inhibiting development of sec. cataracts after  
FT extra:capsular cataract extraction.  
XX  
XX Disclosure; Fig.4; 68pp; English.  
XX  
XX The immunotoxin given in AAR70827 comprises the heavy and light chain  
CC variable regions of anti-lens epithelium IgG3 MAB 4197X linked to  
CC ricin-A and a hexa-histidine tag. The DNA construct encoding the  
CC immunotoxin was expressed from pHBL19 in E. coli.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 554 AA;  
Query Match 97.9%; Score 955; DB 16; Length 554;  
Best Local Similarity 95.0%; Pred. No. 9.9e-94;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 DB 278 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINQRFILV 337  
 QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 111  
 DB 338 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 397  
 QY 112 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTLARSFIICIMISEAAR 171  
 DB 398 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTLARSFIICIMISEAAR 457  
 QY 172 FOYIEGEMRTRIRYNRRA 190  
 DB 458 FOYIEGEMRTRIRYNRRA 476  
 RESULT 9  
 AAP90079  
 ID AAP90079 standard; protein; 562 AA.  
 XX  
 AC AAP90079;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Ricin D.  
 XX  
 KW Ricin D; Ricinus communis; castor beans; Zanicariensis variety;  
 KW modified; lectin binding removed; reduced cell binding  
 XX  
 OS Ricinus communis (castor beans).  
 XX  
 PN W08904839-A.  
 XX  
 PD 01-JUN-1989.  
 XX  
 PF 23-NOV-1988; 88WO-US04238.  
 XX  
 PR 24-NOV-1987; 87US-0124735.  
 XX  
 PA (GEMY ) GENETICS INST INC.  
 XX  
 PI Brown EL, Jones S;  
 XX  
 DR WPI; 1989-178366/24.  
 DR N-PSDB; AAN90068.  
 XX  
 PT Modified ricin molecules and toxin conjugates  
 PT - in which the lectin binding function of the B chain  
 PT is removed or diminished to reduce cell binding.  
 XX  
 PS Disclosure; fig 1; 51pp; English.  
 XX  
 CC Ricin D (see corresp. AAN90068) found on the EcoRI-HindIII fragment  
 CC of DNA from Ricinus communis, Zanicariensis variety. Patent  
 CC discloses many modifications of ricin in which the lectin binding  
 CC function of the B chain is diminished or removed, and conjugation  
 CC to toxins to eliminate cell binding.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 562 AA;  
 Query Match 97.9%; Score 955; DB 10; Length 562;  
 Best Local Similarity 95.0%; Pred. No. 1e-93;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 DB 36 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINQRFILV 95  
 QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 111

DB 96 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 155  
 QY 112 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTLARSFIICIMISEAAR 171  
 DB 156 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTLARSFIICIMISEAAR 215  
 QY 172 FOYIEGEMRTRIRYNRRA 190  
 DB 216 FOYIEGEMRTRIRYNRRA 234  
 RESULT 10  
 AAP50166  
 ID AAP50166 standard; Protein; 565 AA.  
 XX  
 AC AAP50166;  
 XX  
 DT 16-OCT-1991 (first entry)  
 XX  
 DE Sequence of preprorin encoded by pRCL617.  
 XX  
 KW Toxin; anti-tumour therapy.  
 XX  
 OS Ricinus.  
 XX  
 FH Key  
 FT Peptide Location/Qualifiers  
 FT 1..24  
 FT /label= signal  
 FT Protein 25..565  
 FT Region 292..303  
 FT /label= links the C-terminus of the A chain and  
 FT /label= the N-terminus of the B chain  
 FT  
 FT Modified-site 34..36  
 FT /label= N-linked glycosylation  
 FT Modified-site 260..262  
 FT /label= N-linked glycosylation  
 FT Modified-site 398..400  
 FT /label= N-linked glycosylation  
 FT Modified-site 438..440  
 FT /label= N-linked glycosylation  
 FT  
 PN EP145111-A.  
 XX  
 PD 19-JUN-1985.  
 XX  
 PF 13-JUL-1984; 84EP-0304801.  
 XX  
 PR 13-MAR-1984; 84GB-0006569.  
 PR 15-JUL-1983; 83GB-0019265.  
 PR 15-JUL-1983; 83CH-0019265.  
 XX  
 PA (UYWA-) UNIV WARWICK.  
 XX  
 PI Lord JM, Roberts LM, Lamb FI;  
 XX  
 DR WPI; 1985-148040/25.  
 DR N-PSDB; AAN50202.  
 XX  
 PT New DNA sequences coding for ricin type plant toxin - or its  
 PT mutants, and modified vectors and host microorganisms  
 XX  
 PS Disclosure; Page 30-30c; 40pp; English.  
 XX  
 CC Preprorin is the whole polypeptide encoded by AAN50202 and the DNA  
 CC encoding this is claimed. Preprorin is obtained from preprorin by  
 CC removal of the AA leader sequence. The linker AA sequence which is  
 CC present in the precursor polypeptide is enzymatically removed in the  
 CC cell to separate the A and B chains, which are joined by a  
 CC disulphide bridge during the formation of the ricin molecule itself.  
 CC This linker region as well as the presumptive amino terminal leader  
 CC or signal sequence are not present in the sequences already  
 CC published by Funatsu et al.

XX SQ Sequence 565 AA;  
 Query Match 97.9%; Score 955; DB 6; Length 565;  
 Best Local Similarity 95.0%; Pred. No. 1e-93; 0; Indels 10; Gaps 1;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 DB 25 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGADVRHDIPVLPNRVGLPINQRFILV 84  
 QY 52 ELSNHAELS VTLALDVTNAYVGYRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 111  
 DB 85 ELSNHAELS VTLALDVTNAYVGYRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 144  
 QY 112 GNYDRLEQLAGNLRNLELNGPLEEAI SALLYSTGTGTLPTLARSFIICIQMISEAAR 171  
 DB 145 GNYDRLEQLAGNLRNLELNGPLEEAI SALLYSTGTGTLPTLARSFIICIQMISEAAR 204  
 QY 172 FOYIEGEMTRIRYNRRSA 190  
 DB 205 FOYIEGEMTRIRYNRRSA 223  
 RESULT 11  
 AAG78300 standard; Protein; 565 AA.  
 XX ID AAG78300;  
 XX AC AAG78300;  
 XX DT 15-NOV-2001 (first entry)  
 XX DE Castor bean preproricin protein (SEQ ID 1).  
 XX KW Castor bean plant; preproricin; ricin; A chain; B chain;  
 XX KW human immunodeficiency virus infection; HIV; toxin; antiviral agent;  
 XX KW retroviral infection; anti-HIV; virucide activity; viral protease.  
 XX OS Ricinus communis.  
 XX FH Location/Qualifiers  
 XX FT 1..24  
 XX FT /label= Signal peptide  
 XX FT 25..290  
 XX FT /label= Ricin A chain  
 XX FT /note= "N-glycosidase"  
 XX FT 291..302  
 XX FT /label= Linker peptide  
 XX FT /note= "Cleaved during activation of ricin"  
 XX FT 303..565  
 XX FT /label= Ricin B chain  
 XX FT /note= "Galactose/N-acetylglactosamine-binding lectin"  
 XX PN WO200160393-A1.  
 XX PD 23-AUG-2001.  
 XX PF 15-FEB-2001; 2001WO-US05282.  
 XX PR 16-FEB-2000; 2000US-0182759.  
 XX PA (BECH-) BECHTEL BWXT IDAHO LLC.  
 XX PI Keener WK, Ward TE;  
 XX DR WPI; 2001-581908/65.  
 XX DR N-PSDB; AAI64137.  
 XX PT Novel composition comprising toxin e.g., ricin based antiviral compound  
 XX PT useful for treating viral infections such as human immunodeficiency  
 XX PT virus infection.  
 XX PS Disclosure; Page 47-50; 66pp; English.

XX CC The sequence relates to preproricin protein encoded by the DNA sequence  
 CC given in AAI64137. The invention relates to a novel toxin (e.g., ricin)  
 CC based antiviral agent which is toxic to virus-infected cells, but  
 CC non-toxic to uninfected cells. The invention has anti-HIV and virucide  
 CC activities. Its mechanism of action is through inactivation of cellular  
 CC ribosomes and enhancement of binding of the antiviral agent to galactose  
 CC residues on cell surfaces, and its cellular internalisation. The  
 CC invention is useful for treating human immunodeficiency virus infection  
 CC and other viral infections, especially retroviral infections. The  
 CC antiviral agent is activated in viral particles or early-stage infected  
 CC cells, killing the cells upon infection and effectively preventing the  
 CC integration of the viral genome into the host genome thereby preventing  
 CC the latency/rebound problem. The agent enters all HIV susceptible cells,  
 CC and not just cells known to act as host cells for the virus. The  
 CC antiviral agent remains inert in a cell until degraded in it, unless the  
 CC cell is infected with the virus, where the viral protease activates it.  
 XX SQ Sequence 565 AA;  
 Query Match 97.9%; Score 955; DB 22; Length 565;  
 Best Local Similarity 95.0%; Pred. No. 1e-93; 0; Indels 10; Gaps 1;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 DB 25 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGADVRHDIPVLPNRVGLPINQRFILV 84  
 QY 52 ELSNHAELS VTLALDVTNAYVGYRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 111  
 DB 85 ELSNHAELS VTLALDVTNAYVGYRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 144  
 QY 112 GNYDRLEQLAGNLRNLELNGPLEEAI SALLYSTGTGTLPTLARSFIICIQMISEAAR 171  
 DB 145 GNYDRLEQLAGNLRNLELNGPLEEAI SALLYSTGTGTLPTLARSFIICIQMISEAAR 204  
 QY 172 FOYIEGEMTRIRYNRRSA 190  
 DB 205 FOYIEGEMTRIRYNRRSA 223  
 RESULT 12  
 AAG78304 standard; Protein; 565 AA.  
 XX ID AAG78304;  
 XX AC AAG78304;  
 XX DT 27-NOV-2001 (first entry)  
 XX DE Modified castor bean preproricin (SEQ ID 10).  
 XX KW Castor bean plant; preproricin; ricin; A chain; B chain;  
 XX KW human immunodeficiency virus infection; HIV; toxin; antiviral agent;  
 XX KW retroviral infection; anti-HIV; virucide; viral protease.  
 XX OS Chimeric - Ricinus communis  
 XX OS Chimeric - Human immunodeficiency virus type 2.  
 XX FH Location/Qualifiers  
 XX FT 1..24  
 XX FT /label= Signal peptide  
 XX FT 25..565  
 XX FT /label= Proridin  
 XX FT /note= "Proridin consists of the ricin A chain, a linker  
 XX FT peptide, and the ricin B chain. Proridin is  
 XX FT proteolytically cleaved between the A chain and  
 XX FT the linker to yield mature ricin"  
 XX FT 25..291  
 XX FT /label= Ricin A chain  
 XX FT /note= "N-glycosidase"  
 XX FT 292..303  
 XX FT /label= Linker peptide  
 XX FT Cleavage-site 296..297

FT FT /label= HIV protease\_cleavage\_site  
 FT 304..565  
 FT /label= Ricin B chain  
 FT /note= "Galactose/N-acetylgalactosamine-binding lectin"  
 XX  
 XX WO200160393-A1  
 XX  
 XX 23-AUG-2001.  
 XX  
 XX 15-FEB-2001; 2001WO-US05282.  
 XX  
 XX 16-FEB-2000; 2000US-0182759.  
 XX  
 XX (BECH-) BECHTEL BWXT IDAHO LLC.  
 XX  
 XX Keener WK, Ward TE;  
 XX  
 XX WPI: 2001-581908/65.  
 XX N-PSDB; AA164145.  
 XX  
 XX Novel composition comprising toxin e.g., ricin based antiviral compound  
 PT useful for treating viral infections such as human immunodeficiency  
 PT virus infection.  
 PT  
 XX Example 1; Page 59-63; 66pp; English.  
 XX  
 XX The sequence relates to the amino acid sequence of a modified prepropricin  
 CC protein encoded by AA164145. The invention relates to a novel toxin  
 CC (e.g. ricin) based antiviral agent which is toxic to virus-infected  
 CC cells, but non-toxic to uninfected cells. The invention has anti-HIV and  
 CC virucide activities. The agent is able to enter all HIV susceptible  
 CC cells, and not just cells known to act as host cells for the virus. The  
 CC antiviral agent remains inert in a cell unless the cell is infected  
 CC with the HIV virus, where the viral protease activates it. Ricin's  
 CC mechanism of action is through inactivation of cellular ribosomes and  
 CC enhancement of binding of the antiviral agent to galactose residues on  
 CC cell surfaces, and its cellular internalisation. The invention is useful  
 CC for treating human immunodeficiency virus infection and other viral  
 CC infections, especially retroviral infections. The antiviral agent is  
 CC activated in viral particles or early-stage infected cells, killing the  
 CC cells upon infection and effectively preventing the integration of the  
 CC viral genome into the host genome thereby preventing the latency/rebound  
 CC problem.  
 XX  
 XX Sequence 565 AA;  
 SQ  
 Query Match 97.9%; Score 955; DB 22; Length 565;  
 Best Local Similarity 95.0%; Pred. No. 1e-93;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 Db 25 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRLTGDVVRHEIPVLPNRVGLPINQRFILV 84  
 QY 52 ELSNHAELS VTLALDVNTNAVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111  
 Db 85 ELSNHAELS VTLALDVNTNAVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 144  
 QY 112 GNYDRLEQLAGNLRENIELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAR 171  
 Db 145 GNYDRLEQLAGNLRENIELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAR 204  
 QY 172 FOYIEGEMTRIRYNRNSA 190  
 Db 205 FOYIEGEMTRIRYNRNSA 223  
 RESULT 13  
 AAP70326  
 ID AAP70326 standard; Protein; 576 AA.  
 XX  
 XX AAP70326;  
 AC  
 XX

DT 25-MAR-2003 (updated)  
 DT 21-MAY-1991 (first entry)  
 XX  
 XX Sequence of Ricinus communis (castor bean) Ricin toxin  
 DE (RT or ricin) E precursor encoded by PRT38.  
 DE  
 XX Lectin; toxin protein; cytotoxic; cytostatic; castor bean;  
 KW plant toxin.  
 KW  
 XX Ricinus communis.  
 OS  
 XX Location/Qualifiers  
 PH Key 1..35  
 FT Region /note= "leader"  
 FT 36..302  
 FT Region /note= "A-chain"  
 FT 315..576  
 FT Region /note= "B-chain"  
 XX  
 XX EP237676-A.  
 XX  
 XX 23-SEP-1987.  
 PD  
 XX 13-NOV-1986; 86EP-0308877.  
 PF  
 XX 07-MAR-1986; 86US-0837583.  
 PR  
 XX (CETU ) CETUS CORP.  
 PA (CHIR ) CHIRON CORP.  
 XX  
 XX Piatak M;  
 XX  
 XX WPI: 1987-265177/38.  
 DR N-PSDB; AA70526.  
 DR  
 XX New non-glycosylated ricin precursor and toxin etc. - are prepd.  
 PT by recombinant DNA procedures with specific isolation steps for  
 PT purer and soluble prods.  
 PT  
 XX Disclosure; Fig 14(1-2); 112pp; English.  
 PS  
 XX The full length sequences encoding ricin A (AA70520), ricin D  
 CC (AA70525) putative ricin E (AA70526) and RCA (AA70524) in precursor  
 CC form were obtained, using the messenger RNA to obtain a cDNA library, and  
 CC then probing the library to retrieve the desired cDNA inserts. The  
 CC library was probed using the 35-mer given in AA70514. Figure 4 (see  
 CC AA70520, AA70521, AA70522), shows the nucleotide sequences of three  
 CC plasmids containing cDNA inserts obtained by probing a cDNA library  
 CC for sequences encoding ricin B using the probe in AA70517. The cDNA  
 CC inserts can be placed into expression vectors. Site-directed  
 CC mutagenesis may be used to place an ATG start codon and a HindIII  
 CC site at the beginning of the mature protein, (see AA70519). The  
 CC coding sequences of the inserts can be ligated into expression  
 CC vectors containing the PhOA promoter-operator and leader sequence  
 CC (AA70523) and suitable retroregulators.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 XX Sequence 576 AA;  
 SQ  
 Query Match 97.9%; Score 955; DB 8; Length 576;  
 Best Local Similarity 95.0%; Pred. No. 1e-93;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 Db 36 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRLTGDVVRHEIPVLPNRVGLPINQRFILV 95  
 QY 52 ELSNHAELS VTLALDVNTNAVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111  
 Db 96 ELSNHAELS VTLALDVNTNAVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 155  
 QY 112 GNYDRLEQLAGNLRENIELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAR 171

Db 156 GNYDRLEQLAGNLRNENLGNGLPELEAISALYYSTGGTQLPTLARSFFIICQMISEAR 215  
 QY 172 FOYIEGEMTRIRYNRSA 190  
 Db 216 FOYIEGEMTRIRYNRSA 234

RESULT 14  
 AAW25787  
 ID AAW25787 standard; Protein; 576 AA.  
 XX  
 AC AAW25787;  
 DT 25-MAR-2003 (updated)  
 DT 27-MAR-1998 (first entry)  
 XX  
 DE Castor bean ricin.  
 XX  
 KW Ricin; cytotoxin; hybrid protein; cell delivery;  
 KW cell binding ligand; translocation domain; diphtheria toxin B';  
 KW interleukin-2; T-cell lymphoma; organ rejection; therapy.  
 XX  
 OS Ricinus communis.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..35  
 FT /label= Sig\_peptide  
 FT Protein 36..302  
 FT /label= A-domain  
 FT Peptide 303..314  
 FT /label= Linker  
 FT Domain 315..576  
 FT /label= B-domain

US5668255-A.  
 16-SEP-1997.  
 04-AUG-1993; 93US-0102387.  
 27-JUN-1991; 91US-0722484.  
 07-JUN-1984; 84US-0618199.  
 25-APR-1985; 85US-0726808.  
 07-JUN-1985; 85US-0742554.  
 22-DEC-1989; 89US-0456095.  
 14-JUN-1990; 90US-0538276.  
 04-AUG-1993; 93US-0102387.  
 (SERA-) SERAGEN INC.  
 Murphy JR;  
 WPI; 1997-470103/43.  
 N-ESDB; AAT91638.  
 New hybrid molecules for delivery of agents to cells - comprise a binding domain of a cell binding ligand and a portion of a translocation domain of a protein

Example 4; Fig 11A-B; 30pp; English.  
 This polypeptide comprises the castorbean cytotoxin, ricin. DNA (see AAT91638) encoding the enzymatic A domain and a portion of the A-to-B linker peptide of ricin was used to construct a ricin-diphtheria toxin B'-interleukin-2 gene that was expressed in E. coli. The hybrid protein can be isolated and used to treat conditions involving over-production of cells bearing IL2 receptors, such as certain T-cell lymphomas and organ transplant rejection crises. The hybrid inactivates ribosomes in cells bearing IL2 receptors, resulting in cessation of protein synthesis and death of target cells. Claimed hybrid proteins comprise a translocation domain and a cell binding domain from e.g. a hormone, growth factor or polypeptide toxin. The hybrid molecules can be used for the

CC delivery of agents (e.g. therapeutic genes, toxins, detectable labels) into cells. The use of a translocation mechanism ensures that the hybrid will be effective in relatively low doses, since a high proportion of the substance of interest will be taken into the targeted cells. The hybrid molecules can be manufactured as a single hybrid recombinant protein, permitting reproducibility, consistency, and the precise control of composition.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 576 AA;  
 Query Match 97.9%; Score 955; DB 18; Length 576;  
 Best Local Similarity 95.0%; Pred. No. 1e-93; Indels 10; Gaps 1;  
 Matches 189; Conservative 0; Mismatches 0;  
 QY 2 IFPKQYPIINFITAGATVQSYTNFIRAVRGRUT-----VLPNRVGLPINQRFILV 51  
 Db 36 IFPKQYPIINFITAGATVQSYTNFIRAVRGRUT-----VLPNRVGLPINQRFILV 95  
 QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFPHPNQEDAEATHLFTDVQNRYYTFAFG 111  
 Db 96 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFPHPNQEDAEATHLFTDVQNRYYTFAFG 155  
 QY 112 GNYDRLEQLAGNLRNENLGNGLPELEAISALYYSTGGTQLPTLARSFFIICQMISEAR 171  
 Db 156 GNYDRLEQLAGNLRNENLGNGLPELEAISALYYSTGGTQLPTLARSFFIICQMISEAR 215  
 QY 172 FOYIEGEMTRIRYNRSA 190  
 Db 216 FOYIEGEMTRIRYNRSA 234

RESULT 15  
 AAY55892  
 ID AAY55892 standard; Protein; 576 AA.  
 XX  
 AC AAY55892;  
 DT 15-FEB-2000 (first entry)  
 XX  
 DE Castor bean ricin toxin.  
 XX  
 KW Recombinant; hybrid; binding domain; ligand; animal cell; diphtheria; translocation domain; botulinum; neurotoxin; ricin; cholera; tetanus; shiga-like toxin; pertussis; translocation; cytoplasmic membrane; HIV; cytosol; therapy; genetic deficiency disease; enzyme; co-factor; poison; adipocyte; cancer; virus; infection; antibody.  
 XX  
 OS Ricinus communis.  
 XX  
 PN US5965406-A.  
 PD 12-OCT-1999.  
 XX  
 PF 07-JUN-1995; 95US-0488246.  
 XX  
 PR 04-AUG-1993; 93US-0102387.  
 PR 07-JUN-1984; 84US-0618199.  
 PR 27-JUN-1991; 91US-0722484.  
 PR 25-APR-1985; 85US-0726808.  
 PR 07-JUN-1985; 85US-0742554.  
 PR 22-DEC-1989; 89US-0456095.  
 PR 14-JUN-1990; 90US-0538276.  
 XX  
 PA (SERA-) SERAGEN INC.  
 PI Murphy JR;  
 XX  
 DR WPI; 1999-632431/54.  
 DR N-ESDB; AAZ30663.  
 XX  
 PT Recombinant DNA molecule encoding a three part hybrid protein used in the treatment of Aids and genetic deficiency diseases -

XX

Example 4; Fig 11; 31pp; English.

PS The invention relates to a recombinant DNA molecule encoding a hybrid  
XX protein comprising three parts: (a) the first part comprises a portion  
CC of the binding domain of a cell-binding polypeptide ligand allowing the  
CC hybrid protein to bind to an animal cell; (b) the second part comprises  
CC a portion of a translocation domain of a naturally occurring protein  
CC selected from diphtheria toxin, botulinum neurotoxin, ricin, cholera  
CC toxin, LT toxin, C3 toxin, shiga-like toxin, pertussis toxin and tetanus  
CC toxin, which translocate the third part of the across the cytoplasmic  
CC membrane into the cytosol of the cell; and (c) the third part comprises  
CC a polypeptide entity to be introduced into the cell, which is non-native  
CC to the naturally occurring protein of (b). This sequence represents the  
CC Castor bean ricin toxin sequence for use in generating the hybrid of the  
CC invention. The hybrid molecule enables the direction of appropriate  
CC therapy to affected cells, allowing them to function properly and  
CC alleviate or cure the disease. The hybrid is especially used in treating  
CC genetic deficiency diseases, by delivering to affected cells an enzyme  
CC supplying the missing function, to supplementing cellular levels of a  
CC particular enzyme or a scarce precursor or cofactor, to directing toxins  
CC or other poisons to destroy particular cells (such as adipocytes, cancer  
CC cell, or virus infected-cells), to counteracting viral infections such as  
CC HIV, by introducing appropriate antibodies to viral proteins. It is also  
CC involved in the process of getting non-therapeutic substances such as  
CC detectable labels into cells.

XX

SQ Sequence 576 AA;

Query Match 97.9%; Score 955; DB 20; Length 576;  
Best Local Similarity 95.0%; Pred. No. 1e-93; Mismatches 0; Indels 10; Gaps 1;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IPPKQYPIINFATTAGATVQSYNFIKAVRGRLT-----VLENRVGLPINQRFILV 51  
DB 36 IPPKQYPIINFATTAGATVQSYNFIKAVRGRLTGGADVRHEIPVLENRVGLPINQRFILV 95  
QY 52 ELSNHAELSVTLALDVNTAYVVGVRAGNSAYFFHPDNDQEDAEATHLFTDVQNRYYTFAFG 111  
DB 96 ELSNHAELSVTLALDVNTAYVVGVRAGNSAYFFHPDNDQEDAEATHLFTDVQNRYYTFAFG 155  
QY 112 GNYDRLEQLAGNLRNIEELGNGLPEEALISALYYSTGGTGLPTLARSFIIQIMISEAAR 171  
DB 156 GNYDRLEQLAGNLRNIEELGNGLPEEALISALYYSTGGTGLPTLARSFIIQIMISEAAR 215  
QY 172 FOYIEGEMTRIRYNRRA 190  
DB 216 FOYIEGEMTRIRYNRRA 234

Search completed: February 10, 2004, 16:22:29  
Job time : 32.3944 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:18:30 ; Search time 10.9859 Seconds  
(without alignments)  
731.761 Million cell updates/sec

Title: US-10-083-336A-11

Perfect score: 975  
Sequence: 1 MIPKQYPIINFAGATVQ.....RFQYIEGEMTRIRYNRRA 190

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pap.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pap.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pap.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pap.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pap.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	960	98.5	268	2	US-08-356-786-8
2	960	98.5	534	2	US-08-356-786-10
3	955	97.9	267	1	US-07-901-707-1
4	955	97.9	267	1	US-07-988-430-1
5	955	97.9	267	1	US-08-425-336-1
6	955	97.9	267	1	US-08-488-113B-1
7	955	97.9	267	1	US-08-477-484B-1
8	955	97.9	267	2	US-08-646-360-1
9	955	97.9	267	3	US-08-839-765-1
10	955	97.9	267	3	US-09-136-389-1
11	955	97.9	267	4	US-09-610-838-1
12	955	97.9	290	5	PCT-US92-09487-1
13	955	97.9	290	1	US-08-378-761A-27
14	955	97.9	290	1	US-08-485-286-27
15	955	97.9	290	6	5248606-4
16	945	96.9	267	1	US-08-218-303-16
17	945	96.9	267	2	US-08-338-793D-61
18	945	96.9	267	4	US-09-538-873-1
19	865.5	88.8	540	1	US-08-378-761A-77
20	865.5	88.8	540	1	US-08-485-286-77
21	336	34.5	247	1	US-08-488-113B-6
22	336	34.5	247	1	US-08-477-484B-6
23	336	34.5	247	2	US-08-646-360-6
24	336	34.5	247	3	US-08-839-765-6
25	336	34.5	247	3	US-09-136-389-6
26	336	34.5	247	4	US-09-610-838-6
27	336	34.5	267	1	US-08-378-761A-74

28	336	34.5	267	1	US-08-485-286-74	Sequence 74, Appl
29	336	34.5	289	1	US-07-923-692C-4	Sequence 4, Appl
30	336	34.5	289	1	US-08-184-237-4	Sequence 4, Appl
31	336	34.5	289	2	US-08-482-920-4	Sequence 4, Appl
32	336	34.5	289	3	US-08-484-341-4	Sequence 4, Appl
33	336	34.5	289	3	US-08-483-502-4	Sequence 4, Appl
34	336	34.5	289	4	US-09-726-651A-4	Sequence 4, Appl
35	320.5	32.9	282	1	US-08-324-301-15	Sequence 15, Appl
36	310.5	31.8	255	1	US-07-901-707-6	Sequence 6, Appl
37	310.5	31.8	255	1	US-07-988-430-6	Sequence 6, Appl
38	310.5	31.8	255	1	US-08-425-336-6	Sequence 6, Appl
39	310.5	31.8	255	5	PCT-US92-09487-6	Sequence 7, Appl
40	305	31.3	248	3	US-08-902-486-7	Sequence 2, Appl
41	305	31.3	290	1	US-08-245-754A-2	Sequence 2, Appl
42	305	31.3	290	2	US-08-597-731-2	Sequence 15, Appl
43	305	31.3	496	3	US-08-902-486-15	Sequence 71, Appl
44	303.5	31.1	250	1	US-08-378-761A-71	Sequence 71, Appl
45	303.5	31.1	250	1	US-08-485-286-71	Sequence 71, Appl

ALIGNMENTS

RESULT 1

US-08-356-786-8

; Sequence 8, Application US/08356786

; Patent No. 5877305

; GENERAL INFORMATION:

; APPLICANT: Huston, James S.

; APPLICANT: Oppermann, Hermann

; APPLICANT: Houston, L. L.

; APPLICANT: Ring, David B.

; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer

; TITLE OF INVENTION: Marker

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault

; STREET: Exchange Place, 53 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/356,786

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/831,967

; FILING DATE: 06-FEB-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Pitcher, Edmund R.

; REGISTRATION NUMBER: 27,829

; REFERENCE/DOCKET NUMBER: CRP-053

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 248-7000

; TELEFAX: (617) 248-7100

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 268 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-356-786-8

Query Match 98.5%; Score 960; DB 2; Length 268;

Best Local Similarity 95.0%; Pred. No. 2.5e-103;

Matches 190; Conservative 0; Mismatches 0; Indels 10; Gaps 1;



QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 50  
 Db 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFIL 60  
 QY 51 VELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDNQDEAEATHLFTDVQNRYYTFAF 110  
 Db 61 VELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDNQDEAEATHLFTDVQNRYYTFAF 120  
 QY 111 GGNVDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFFIICQMISEAA 170  
 Db 121 GGNVDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFFIICQMISEAA 180  
 QY 171 RFOYIEGEMRTRIRYNRRA 190  
 Db 181 RFOYIEGEMRTRIRYNRRA 200

## RESULT 2

US-08-356-786-10  
 ; Sequence 10, Application US/08356786  
 ; Patent No. 587305  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Huston, James S.  
 ; APPLICANT: Oppermann, Hermann  
 ; APPLICANT: Houston, L. L.  
 ; APPLICANT: King, David B.  
 ; TITLE OF INVENTION: Bicosynthetic Binding Protein for Cancer  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
 ; STREET: Exchange Place, 53 State Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/356,786  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/831,967  
 FILING DATE: 06-FEB-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pitcher, Edmund R.  
 REGISTRATION NUMBER: 27,829  
 REFERENCE/DOCKET NUMBER: CRP-053  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 248-7000  
 TELEFAX: (617) 248-7100  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 534 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-356-786-10

Query Match 98.5%; Score 960; DB 2; Length 534;  
 Best Local Similarity 95.0%; Pred. No. 6.8e-103;  
 Matches 190; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 50  
 Db 3 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFIL 62  
 QY 51 VELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDNQDEAEATHLFTDVQNRYYTFAF 110

Db 63 VELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDNQDEAEATHLFTDVQNRYYTFAF 122  
 QY 111 GGNVDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFFIICQMISEAA 170  
 Db 123 GGNVDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFFIICQMISEAA 182  
 QY 171 RFOYIEGEMRTRIRYNRRA 190  
 Db 183 RFOYIEGEMRTRIRYNRRA 202

## RESULT 3

US-07-901-707-1  
 ; Sequence 1, Application US/07901707  
 ; Patent No. 5376546  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bernhardt, Susan L.  
 ; APPLICANT: Better, Marc D.  
 ; APPLICANT: Carroll, Steve F.  
 ; APPLICANT: Lane, Julie A.  
 ; TITLE OF INVENTION: Materials Comprising and Methods of  
 ; TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins  
 ; NUMBER OF SEQUENCES: 57  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
 ; ADDRESSEE: Bicknell  
 ; STREET: Two First National Plaza, 20 South Clark  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60603  
 ; COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/901,707  
 FILING DATE: 19920619  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/787,567  
 FILING DATE: 04-NOV-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: No. 5376546and, Greta E.  
 REGISTRATION NUMBER: 35,302  
 REFERENCE/DOCKET NUMBER: 27129/30910  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (312) 346-5750  
 TELEFAX: (312) 984-5750  
 TELEX: 25-3856  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 267 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-07-901-707-1

Query Match 97.9%; Score 955; DB 1; Length 267;  
 Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 51  
 Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFIL 60  
 QY 52 VELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDNQDEAEATHLFTDVQNRYYTFAF 111  
 Db 61 VELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDNQDEAEATHLFTDVQNRYYTFAF 120  
 QY 112 GGNVDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFFIICQMISEAA 171

Db 121 GNYDRLEQLAGNLRENIELGNGLPEEAISALYYSTGGTQLPTLARSFICQMISEAAR 180  
QY 172 FOYIEGEMTRIRYNRSA 190  
Db 181 FOYIEGEMTRIRYNRSA 199

## RESULT 4

US-07-988-430-1  
; Sequence 1, Application US/07988430  
; Patent No. 5416202  
; GENERAL INFORMATION:  
; APPLICANT: Bernhard, Susan L.  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Lane, Julie A.  
; APPLICANT: Lei, Shau-Ping  
; TITLE OF INVENTION: Materials Comprising and Methods of  
; Preparation and Use for Ribosome-Inactivating Proteins  
; NUMBER OF SEQUENCES: 101  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/988,430  
; FILING DATE: 19921209  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5416202and, Greta E.  
; REGISTRATION NUMBER: 35302  
; REFERENCE/DOCKET NUMBER: 31133  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-988-430-1

Query Match 97.9%; Score 955; DB 1; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNRYYTFAFG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNRYYTFAFG 120

QY 112 GNYDRLEQLAGNLRENIELGNGLPEEAISALYYSTGGTQLPTLARSFICQMISEAAR 171  
Db 121 GNYDRLEQLAGNLRENIELGNGLPEEAISALYYSTGGTQLPTLARSFICQMISEAAR 180  
QY 172 FOYIEGEMTRIRYNRSA 190  
Db 181 FOYIEGEMTRIRYNRSA 199

## RESULT 5

US-08-425-336-1  
; Sequence 1, Application US/08425336  
; Patent No. 5621083  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; Proteins  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/425,336  
; FILING DATE: 18-APR-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/064,691  
; FILING DATE: 12-MAY-1993  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Thomas C.  
; REGISTRATION NUMBER: P-36,989  
; REFERENCE/DOCKET NUMBER: 31394  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-425-336-1

Query Match 97.9%; Score 955; DB 1; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNRYYTFAFG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNRYYTFAFG 120  
QY 112 GNYDRLEQLAGNLRENIELGNGLPEEAISALYYSTGGTQLPTLARSFICQMISEAAR 171

Db 121 GNYDRLEQAGNLRNIELGNGPLEEASALYYSTGGTQLPTLARSFIIQIMISEAAR 180  
QY 172 FOYIEGEMTRIRYNRRA 190  
Db 181 FOYIEGEMTRIRYNRRA 199  
RESULT 6  
US-08-488-113B-1  
; Sequence 1, Application US/08488113B  
; Patent No. 5744580  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,113B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/425,336  
; FILING DATE: 18-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-488-113B-1  
Query Match 97.9%; Score 955; DB 1; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQVPIINFTTAGTQVSYTTFIRAVRGRLT-----VLNRYVGLPINORFILV 51  
Db 1 IFPKQVPIINFTTAGTQVSYTTFIRAVRGRLTGTGADVRIEIPVLENRVGLPINORFILV 60

QY 52 ELSNHAELSVTIALDVTNAYVYGRAGNSAYFFHPDNOEDAEATHLFTDVQNRVTFAG 111  
Db 61 ELSNHAELSVTIALDVTNAYVYGRAGNSAYFFHPDNOEDAEATHLFTDVQNRVTFAG 120  
QY 112 GNYDRLEQAGNLRNIELGNGPLEEASALYYSTGGTQLPTLARSFIIQIMISEAAR 171  
Db 121 GNYDRLEQAGNLRNIELGNGPLEEASALYYSTGGTQLPTLARSFIIQIMISEAAR 180  
QY 172 FOYIEGEMTRIRYNRRA 190  
Db 181 FOYIEGEMTRIRYNRRA 199  
RESULT 7  
US-08-477-484B-1  
; Sequence 1, Application US/08477484B  
; Patent No. 5756699  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,484B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/425,336  
; FILING DATE: 18-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-477-484B-1  
Query Match 97.9%; Score 955; DB 1; Length 267;

Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
Qy 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVREHPIVLPNRVGLPINQRFILV 60  
Qy 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTTAFG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTTAFG 120  
Qy 112 GNYDRLEQLAGNLRNIELGNGLPLEEALISALYYSTGGTQTLPTLARSFIICMISEAAAR 171  
Db 121 GNYDRLEQLAGNLRNIELGNGLPLEEALISALYYSTGGTQTLPTLARSFIICMISEAAAR 180  
Qy 172 FOYIEGEMRTRIYNNRRA 190  
Db 181 FOYIEGEMRTRIYNNRRA 199  
RESULT 8  
US-08-646-360-1  
; Sequence 1, Application US/08646360  
; Patent No. 5837491  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 173  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/646,360  
; FILING DATE: 13-MAY-1996  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/05348  
; FILING DATE: 12-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 200-70.P4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/707-9889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-646-360-1  
Query Match 97.9%; Score 955; DB 2; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
Qy 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVREHPIVLPNRVGLPINQRFILV 60  
Qy 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTTAFG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTTAFG 120  
Qy 112 GNYDRLEQLAGNLRNIELGNGLPLEEALISALYYSTGGTQTLPTLARSFIICMISEAAAR 171  
Db 121 GNYDRLEQLAGNLRNIELGNGLPLEEALISALYYSTGGTQTLPTLARSFIICMISEAAAR 180  
Qy 172 FOYIEGEMRTRIYNNRRA 190  
Db 181 FOYIEGEMRTRIYNNRRA 199  
RESULT 9  
US-08-839-765-1  
; Sequence 1, Application US/08839765  
; Patent No. 6145631  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/839,765  
; FILING DATE: 15-APR-1997  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/425,336  
; FILING DATE: 18-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-839-765-1

Query Match 97.9%; Score 955; DB 3; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLTGDVVRHEIPVLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 120  
QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEAI SALYYSTGGTQLPTLARSFIIQMISEAR 171  
Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEAI SALYYSTGGTQLPTLARSFIIQMISEAR 180  
QY 172 FOYIEGEMTRIRYNRSA 190  
Db 181 FOYIEGEMTRIRYNRSA 199

## RESULT 10

US-09-136-389-1  
Sequence 1, Application US/09136389  
Patent No. 6146850  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,389  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-136-389-1

Query Match 97.9%; Score 955; DB 3; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLTGDVVRHEIPVLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 120  
QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEAI SALYYSTGGTQLPTLARSFIIQMISEAR 171  
Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEAI SALYYSTGGTQLPTLARSFIIQMISEAR 180  
QY 172 FOYIEGEMTRIRYNRSA 190  
Db 181 FOYIEGEMTRIRYNRSA 199

## RESULT 11

US-09-610-838-1  
Sequence 1, Application US/09610838  
Patent No. 6376217  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/610,838  
FILING DATE: 06-JUL-2000  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,389  
FILING DATE: 18-AUG-1998  
APPLICATION NUMBER: 08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348

; FILING DATE: 12-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 200-70.P4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-610-838-1

Query Match 97.9%; Score 955; DB 4; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGDVHRHPIVLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEALTHLFTDVQNYTFAFG 111  
DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEALTHLFTDVQNYTFAFG 120  
QY 112 GNYDRLEQLAGNLRNIELGNGLPLEEALISALYYSTGTGTLPTLARSFIIICQMISEAAR 171  
DB 121 GNYDRLEQLAGNLRNIELGNGLPLEEALISALYYSTGTGTLPTLARSFIIICQMISEAAR 180  
QY 172 FOYIEGEMRTRIRYNRRA 190  
DB 181 FOYIEGEMRTRIRYNRRA 199

RESULT 12  
PCT-US92-09487-1  
; Sequence 1, Application PC/TUS9209487  
; GENERAL INFORMATION:  
; APPLICANT: Bernhard, Susan L.  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Lane, Julie A.  
; APPLICANT: Lei, Shau-Ping  
; TITLE OF INVENTION: Materials Comprising and Methods of  
; Preparation and Use for Ribosome-Inactivating Proteins  
; NUMBER OF SEQUENCES: 101  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/09487  
; FILING DATE: 19921104  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Noland, Greta E.  
; REGISTRATION NUMBER: 35302  
; REFERENCE/DOCKET NUMBER: 31133  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US92-09487-1

Query Match 97.9%; Score 955; DB 5; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGDVHRHPIVLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEALTHLFTDVQNYTFAFG 111  
DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEALTHLFTDVQNYTFAFG 120  
QY 112 GNYDRLEQLAGNLRNIELGNGLPLEEALISALYYSTGTGTLPTLARSFIIICQMISEAAR 171  
DB 121 GNYDRLEQLAGNLRNIELGNGLPLEEALISALYYSTGTGTLPTLARSFIIICQMISEAAR 180  
QY 172 FOYIEGEMRTRIRYNRRA 190  
DB 181 FOYIEGEMRTRIRYNRRA 199

RESULT 13  
US-08-378-761A-27  
; Sequence 27, Application US/08378761A  
; Patent No. 5635384  
; GENERAL INFORMATION:  
; APPLICANT: WALSH, TERENCE A  
; APPLICANT: HEY, TIMOTHY D  
; APPLICANT: MORGAN, ALICE ER  
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
; PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
; USING  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ANDREA T. BORUCKI  
; STREET: 9330 ZIONSVILLE ROAD  
; CITY: INDIANAPOLIS  
; STATE: IN  
; COUNTRY: US  
; ZIP: 46268  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/378,761A  
FILING DATE: 26-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELEPHONE: (317) 337-4846  
TELECOMMUNICATION INFORMATION:  
INFORMATION FOR SEQ ID NO: 27:  
LENGTH: 290 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-378-761A-27

Query Match 97.9%; Score 955; DB 1; Length 290;  
Best Local Similarity 95.0%; Pred. No. 1.1e-102;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
  
QY 2 IPPKQYPIINFTTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51  
Db 25 IPPKQYPIINFTTAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 84  
  
QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAI THLFTDVQNRVTFAG 111  
Db 85 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAI THLFTDVQNRVTFAG 144  
  
QY 112 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQLPTLARSFICIMISEAAR 171  
Db 145 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQLPTLARSFICIMISEAAR 204  
  
QY 172 FOYIEGEMTRIRYNRRA 190  
Db 205 FOYIEGEMTRIRYNRRA 223

RESULT 14  
US-08-485-286-27  
Sequence 27, Application US/08485286  
Patent No. 5646026  
Patent No. 5646026 5646119  
GENERAL INFORMATION:  
APPLICANT: WALSH, TERENCE A  
APPLICANT: HEY, TIMOTHY D  
APPLICANT: MORGAN, ALICE ER  
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
NUMBER OF SEQUENCES: 81  
TITLE OF INVENTION: USING  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDREA T. BORUCKI  
STREET: 9330 ZIONSVILLE ROAD  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,286  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/378761  
FILING DATE: 26-JAN-1995

ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 27:  
LENGTH: 290 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-286-27

Query Match 97.9%; Score 955; DB 1; Length 290;  
Best Local Similarity 95.0%; Pred. No. 1.1e-102;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
  
QY 2 IPPKQYPIINFTTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51  
Db 25 IPPKQYPIINFTTAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 84  
  
QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAI THLFTDVQNRVTFAG 111  
Db 85 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAI THLFTDVQNRVTFAG 144  
  
QY 112 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQLPTLARSFICIMISEAAR 171  
Db 145 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQLPTLARSFICIMISEAAR 204  
  
QY 172 FOYIEGEMTRIRYNRRA 190  
Db 205 FOYIEGEMTRIRYNRRA 223

RESULT 15  
5248606-4  
Patent No. 5248606  
APPLICANT: WALSH, TERENCE A.; HEY, TIMOTHY D.; MORGAN,  
ALICE E.R.  
TITLE OF INVENTION: DNA ENCODING INACTIVE PRECURSOR AND  
ACTIVE FORMS OF MAIZE RIBOSOME INACTIVATING  
NUMBER OF SEQUENCES: 49  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/535,636  
FILING DATE: 11-JUN-1990  
SEQ ID NO: 4:  
LENGTH: 290  
5248606-4

Query Match 97.9%; Score 955; DB 6; Length 290;  
Best Local Similarity 95.0%; Pred. No. 1.1e-102;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
  
QY 2 IPPKQYPIINFTTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51  
Db 25 IPPKQYPIINFTTAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 84  
  
QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAI THLFTDVQNRVTFAG 111  
Db 85 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAI THLFTDVQNRVTFAG 144  
  
QY 112 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQLPTLARSFICIMISEAAR 171  
Db 145 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQLPTLARSFICIMISEAAR 204  
  
QY 172 FOYIEGEMTRIRYNRRA 190  
Db 205 FOYIEGEMTRIRYNRRA 223

Search completed: February 10, 2004, 16:29:35

Job time : 10.9859 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:26:46 ; Search time 24.507 Seconds  
(without alignments)  
1623.314 Million cell updates/sec

Title: US-10-083-336A-11

Perfect score: 975

Sequence: 1 MIFPKQYPIINFITAGATVQ.....RFQYIEGEMRTRINRNSA 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	975	100.0	190	12	US-10-083-336A-11
2	971	99.6	189	12	US-10-083-336A-6
3	966	99.1	188	12	US-10-083-336A-4
4	960	98.5	200	12	US-10-083-336A-10
5	956	98.1	188	12	US-10-083-336A-8
6	956	98.1	199	12	US-10-083-336A-5
7	955	97.9	267	12	US-10-127-890-1
8	955	97.9	576	12	US-10-083-336A-1
9	951	97.9	198	12	US-10-083-336A-3
10	945	96.9	267	12	US-10-282-935-1
11	945	96.9	267	12	US-10-440-796-1
12	941	96.5	198	12	US-10-083-336A-7
13	939.5	96.4	185	12	US-10-083-336A-9
14	679	69.6	179	12	US-10-083-336A-2
15	336	34.5	247	10	US-09-792-793A-39

16	336	34.5	247	12	US-10-127-890-6	Sequence 6, Appli
17	336	34.5	247	12	US-10-375-209A-39	Sequence 39, Appli
18	336	34.5	289	12	US-10-280-679B-4	Sequence 4, Appli
19	305	31.3	247	10	US-09-792-793A-34	Sequence 34, Appli
20	305	31.3	247	12	US-10-375-209A-34	Sequence 34, Appli
21	297.5	30.5	251	12	US-10-282-935-3	Sequence 3, Appli
22	297.5	30.5	251	12	US-10-440-796-3	Sequence 3, Appli
23	273	28.0	263	12	US-10-127-890-7	Sequence 7, Appli
24	272	27.9	263	12	US-10-127-890-4	Sequence 4, Appli
25	257.5	26.4	248	12	US-10-127-890-5	Sequence 5, Appli
26	254	26.1	252	9	US-09-347-064-2	Sequence 2, Appli
27	254	26.1	252	9	US-09-347-064-8	Sequence 8, Appli
28	248.5	25.5	251	12	US-10-127-890-107	Sequence 107, App
29	247.5	25.4	251	12	US-10-127-890-106	Sequence 106, App
30	247.5	25.4	251	12	US-10-127-890-110	Sequence 110, App
31	247.5	25.4	251	12	US-10-127-890-111	Sequence 111, App
32	246.5	25.3	251	9	US-09-765-527-247	Sequence 27, App
33	246.5	25.3	251	12	US-10-127-890-2	Sequence 2, Appli
34	246.5	25.3	251	12	US-10-127-890-99	Sequence 99, Appli
35	246.5	25.3	251	12	US-10-127-890-100	Sequence 100, App
36	246.5	25.3	251	12	US-10-127-890-101	Sequence 101, App
37	246.5	25.3	251	12	US-10-127-890-102	Sequence 102, App
38	246.5	25.3	251	12	US-10-127-890-103	Sequence 103, App
39	246.5	25.3	251	12	US-10-127-890-104	Sequence 104, App
40	246.5	25.3	251	12	US-10-127-890-105	Sequence 105, App
41	246.5	25.3	316	12	US-10-074-596-1	Sequence 1, Appli
42	246.5	25.3	507	12	US-10-074-596-11	Sequence 11, Appli
43	245.5	25.2	293	9	US-09-765-527-259	Sequence 259, App
44	245.5	25.2	309	9	US-09-765-527-253	Sequence 253, App
45	245.5	25.2	332	9	US-09-765-527-251	Sequence 251, App

ALIGNMENTS

RESULT 1

US-10-083-336A-11  
; Sequence 11, Application US/10083336A  
; Publication No. US20030181665A1  
; GENERAL INFORMATION:  
; APPLICANT: Olson, Mark A  
; APPLICANT: Millard, Charles B  
; APPLICANT: Byrne, Michael P  
; APPLICANT: Wannemacher, Robert W  
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof  
; FILE REFERENCE: P67452USO (RiID 01-58)  
; CURRENT APPLICATION NUMBER: US/10/083,336A  
; CURRENT FILING DATE: 2002-05-21  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; TYPE: PRT  
; ORGANISM: Ricinus communis  
US-10-083-336A-11

Query Match	100.0%	Score	975	DB	12	Length	190
Best Local Similarity	100.0%	Pred. No.	1.1e-103				
Matches	190	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MIFPKQYPIINFITAGATVQSYTNFRAVGRITVLPNRVGLPINQRFILVELSNHAELS	60				
Db	1	MIFPKQYPIINFITAGATVQSYTNFRAVGRITVLPNRVGLPINQRFILVELSNHAELS	60				
Qy	61	VTALDVTNAYVGVYRAGNSAYFFHPDNDQDAEAIHTLFTDVQNRYYTFAFGNYDRLEQL	120				
Db	61	VTALDVTNAYVGVYRAGNSAYFFHPDNDQDAEAIHTLFTDVQNRYYTFAFGNYDRLEQL	120				
Qy	121	AGNLRENIELNGPLEEAISALYYSTGQTQLTLARSFFICICIMISEAARFQYIEGEMR	180				
Db	121	AGNLRENIELNGPLEEAISALYYSTGQTQLTLARSFFICICIMISEAARFQYIEGEMR	180				
Qy	181	TRIRYNRNSA	190				

```

Db      181 TRIRYNRSA 190
|||||
RESULT 2
US-10-083-336A-6
; Sequence 6, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-6

Query Match      99.6%; Score 971; DB 12; Length 189;
Best Local Similarity 100.0%; Pred. No. 3.1e-103;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTVLPNRVGLPINQRFILVELSNHAELSV 60
|||||
Db 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTVLPNRVGLPINQRFILVELSNHAELSV 60
|||||
QY 61 VTALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRVTFAPGGNYDRLEQL 120
|||||
Db 61 VTALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRVTFAPGGNYDRLEQL 120
|||||
QY 121 AGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFIIQMISEARFOYIEGEMR 180
|||||
Db 121 AGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFIIQMISEARFOYIEGEMR 180
|||||
QY 181 TRIRYNRS 189
|||||
Db 181 TRIRYNRS 189
|||||

RESULT 3
US-10-083-336A-4
; Sequence 4, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-4

Query Match      99.1%; Score 966; DB 12; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.1e-102;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTVLPNRVGLPINQRFILVELSNHAELSV 61
|||||
Db 181 TRIRYNRS 190
|||||
RESULT 4
US-10-083-336A-10
; Sequence 10, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-10

Query Match      98.5%; Score 960; DB 12; Length 200;
Best Local Similarity 95.0%; Pred. No. 6.1e-102;
Matches 190; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTVLPNRVGLPINQRFIL 50
|||||
Db 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTVLPNRVGLPINQRFIL 60
|||||
QY 51 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRVTFAP 110
|||||
Db 61 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRVTFAP 120
|||||
QY 111 GGNVDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFIIQMISEAA 170
|||||
Db 121 GGNVDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFIIQMISEAA 180
|||||
QY 171 RFQYIEGEMRTRIRYNRSA 190
|||||
Db 181 RFQYIEGEMRTRIRYNRSA 200
|||||

RESULT 5
US-10-083-336A-8
; Sequence 8, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8

```

```

; LENGTH: 188
; TYPE: PRT
; ORGANISM: Ricinus communis
; US-10-083-336A-8

Query Match      98.1%; Score 956; DB 12; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.6e-101;
Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PKQYPIINFNTAGATVQSYTNFIRAVRGRLTVLPNVRVGLPINQRFILVELSNHAELSVTL 63
Db 3 PKQYPIINFNTAGATVQSYTNFIRAVRGRLTVLPNVRVGLPINQRFILVELSNHAELSVTL 62

QY 64 ALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHLFTDQVQRYTFAGGNYDLEQLAGN 123
Db 63 ALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHLFTDQVQRYTFAGGNYDLEQLAGN 122

QY 124 LRENIELGNGLPBEAISALYYSTGGTQPLTLARSFFIICQMISEARFOYIEGEMTRI 183
Db 123 LRENIELGNGLPBEAISALYYSTGGTQPLTLARSFFIICQMISEARFOYIEGEMTRI 182

QY 184 RYNRRS 189
Db 183 RYNRRS 188

RESULT 6
US-10-083-336A-5
; Sequence 5, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 5
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Ricinus communis
; US-10-083-336A-5

Query Match      98.1%; Score 956; DB 12; Length 199;
Best Local Similarity 95.0%; Pred. No. 1.7e-101;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MIFPKQYPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNVRVGLPINQRFIL 50
Db 1 MIFPKQYPIINFNTAGATVQSYTNFIRAVRGRLTGDVVRHEIPVLPNVRVGLPINQRFIL 60

QY 51 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHLFTDQVQRYTF 110
Db 61 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHLFTDQVQRYTF 120

QY 111 GNYDRLEQLAGNLRENIELGNGLPBEAISALYYSTGGTQPLTLARSFFIICQMISEA 170
Db 121 GNYDRLEQLAGNLRENIELGNGLPBEAISALYYSTGGTQPLTLARSFFIICQMISEA 180

QY 171 RFOYIEGEMTRIRYNRRS 189
Db 181 RFOYIEGEMTRIRYNRRS 199

RESULT 7
US-10-127-890-1
; Sequence 1, Application US/10127890
; Publication No. US20030166196A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Better, Marc D.
; Carroll, Stephen F.
; Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/127,890
; FILING DATE: 23-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
; US-10-127-890-1

Query Match      97.9%; Score 955; DB 12; Length 267;
Best Local Similarity 95.0%; Pred. No. 3.5e-101;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNVRVGLPINQRFILV 51
Db 1 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRLTGDVVRHEIPVLPNVRVGLPINQRFILV 60

QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHLFTDQVQRYTF 111
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHLFTDQVQRYTF 120

QY 112 GNYDRLEQLAGNLRENIELGNGLPBEAISALYYSTGGTQPLTLARSFFIICQMISEA 171
Db 121 GNYDRLEQLAGNLRENIELGNGLPBEAISALYYSTGGTQPLTLARSFFIICQMISEA 180

QY 172 RFOYIEGEMTRIRYNRRS 190
Db 181 RFOYIEGEMTRIRYNRRS 199

```

## RESULT 8

US-10-083-336A-1

; Sequence 1, Application US/10083336A  
 ; Publication No. US20030181665A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Olson, Mark A  
 ; APPLICANT: Millard, Charles B  
 ; APPLICANT: Byrne, Michael P  
 ; APPLICANT: Wannemacher, Robert W  
 ; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof  
 ; FILE REFERENCE: P67452US0 (RIID 01-58)  
 ; CURRENT APPLICATION NUMBER: US/10/083,336A  
 ; CURRENT FILING DATE: 2002-05-21  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 576  
 ; TYPE: PRT  
 ; ORGANISM: Ricinus communis  
 US-10-083-336A-1

Query Match 97.9%; Score 955; DB 12; Length 576;  
 Best Local Similarity 95.0%; Pred. No. 1e-100;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQVPIINFITAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 DB 36 IFPKQVPIINFITAGATVQSYTNFIRAVRGRLTGGADVREHIFVLPNRVGLPINQRFILV 95  
 QY 52 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111  
 DB 96 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 155  
 QY 112 GNYDRLEQLAGNLENIELGNGLPLEEALISALYYSTGGTQPLTLARSFFIICQMISEAAR 171  
 DB 156 GNYDRLEQLAGNLENIELGNGLPLEEALISALYYSTGGTQPLTLARSFFIICQMISEAAR 215  
 QY 172 FOYIEGEMTRIRYNNRSA 190  
 DB 216 FOYIEGEMTRIRYNNRSA 234

## RESULT 9

US-10-083-336A-3

; Sequence 3, Application US/10083336A  
 ; Publication No. US20030181665A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Olson, Mark A  
 ; APPLICANT: Millard, Charles B  
 ; APPLICANT: Byrne, Michael P  
 ; APPLICANT: Wannemacher, Robert W  
 ; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof  
 ; FILE REFERENCE: P67452US0 (RIID 01-58)  
 ; CURRENT APPLICATION NUMBER: US/10/083,336A  
 ; CURRENT FILING DATE: 2002-05-21  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 198  
 ; TYPE: PRT  
 ; ORGANISM: Ricinus communis  
 US-10-083-336A-3

Query Match 97.5%; Score 951; DB 12; Length 198;  
 Best Local Similarity 94.9%; Pred. No. 6.5e-101;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQVPIINFITAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 DB 1 IFPKQVPIINFITAGATVQSYTNFIRAVRGRLTGGADVREHIFVLPNRVGLPINQRFILV 60  
 QY 52 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111

DB 61 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 120  
 QY 112 GNYDRLEQLAGNLENIELGNGLPLEEALISALYYSTGGTQPLTLARSFFIICQMISEAAR 171  
 DB 121 GNYDRLEQLAGNLENIELGNGLPLEEALISALYYSTGGTQPLTLARSFFIICQMISEAAR 180  
 QY 172 FOYIEGEMTRIRYNNRSA 189  
 DB 181 FOYIEGEMTRIRYNNRSA 198

## RESULT 10

US-10-282-935-1

; Sequence 1, Application US/10282935  
 ; Publication No. US20030143193A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VITETTA, ELLEN S.  
 ; APPLICANT: GHETTIE, VICTOR F.  
 ; APPLICANT: SMALLSHAW, JOAN G.  
 ; APPLICANT: BALUNA, ROXANA G.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF  
 ; FILE REFERENCE: UTSD:884US  
 ; CURRENT APPLICATION NUMBER: US/10/282,935  
 ; CURRENT FILING DATE: 2002-10-29  
 ; PRIOR APPLICATION NUMBER: 09/538,873  
 ; PRIOR FILING DATE: 2000-03-30  
 ; PRIOR APPLICATION NUMBER: 60/126,826  
 ; PRIOR FILING DATE: 1999-03-30  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 267  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 ; OTHER INFORMATION: Peptide  
 US-10-282-935-1

Query Match 96.9%; Score 945; DB 12; Length 267;  
 Best Local Similarity 94.9%; Pred. No. 4.9e-100;  
 Matches 187; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 4 PKQYPIINFITAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 53  
 DB 3 PKQYPIINFITAGATVQSYTNFIRAVRGRLTGGADVREHIFVLPNRVGLPINQRFILV 62  
 QY 54 SNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 113  
 DB 63 SNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 122  
 QY 114 YDRLEQLAGNLENIELGNGLPLEEALISALYYSTGGTQPLTLARSFFIICQMISEAARFQ 173  
 DB 123 YDRLEQLAGNLENIELGNGLPLEEALISALYYSTGGTQPLTLARSFFIICQMISEAARFQ 182  
 QY 174 YIEGEMTRIRYNNRSA 190  
 DB 183 YIEGEMTRIRYNNRSA 199

## RESULT 11

US-10-440-796-1

; Sequence 1, Application US/10440796  
 ; Publication No. US20040009148A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VITETTA, ELLEN S.  
 ; APPLICANT: GHETTIE, VICTOR F.  
 ; APPLICANT: SMALLSHAW, JOAN G.  
 ; APPLICANT: BALUNA, ROXANA G.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK  
 ; FILE REFERENCE: UTSD:603

CURRENT APPLICATION NUMBER: US/10/440,796  
CURRENT FILING DATE: 2003-05-19  
PRIOR APPLICATION NUMBER: US/09/538,873  
PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: 60/126,826  
PRIOR FILING DATE: 1999-03-30  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 267  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-440-796-1

Query Match 96.9%; Score 945; DB 12; Length 267;  
Best Local Similarity 94.9%; Pred. No. 4.9e-100;  
Matches 187; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 4 PKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILVEL 53  
DB 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRIEIPVLPNRVGLPINQRFILVEL 62  
QY 54 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAGGN 113  
DB 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAGGN 122  
QY 114 YDRLEQLAGNREINELGNPLEEALSALYYSTGTQTLPTLARSFIIICQMISEAARFQ 173  
DB 123 YDRLEQLAGNREINELGNPLEEALSALYYSTGTQTLPTLARSFIIICQMISEAARFQ 182  
QY 174 YIEGEMTRIRYNRS 190  
DB 183 YIEGEMTRIRYNRS 199

RESULT 12  
US-10-083-336A-7  
Sequence 7, Application US/10083336A  
Publication No. US20030181665A1  
GENERAL INFORMATION:  
APPLICANT: Olson, Mark A  
APPLICANT: Millard, Charles B  
APPLICANT: Byrne, Michael P  
APPLICANT: Wannemacher, Robert W  
TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof  
FILE REFERENCE: P67452US0 (RIID 01-58)  
CURRENT APPLICATION NUMBER: US/10/083,336A  
CURRENT FILING DATE: 2002-05-21  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 198  
TYPE: PRT  
ORGANISM: Ricinus communis  
US-10-083-336A-7

Query Match 96.5%; Score 941; DB 12; Length 198;  
Best Local Similarity 94.9%; Pred. No. 9.2e-100;  
Matches 186; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 4 PKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILVEL 53  
DB 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRIEIPVLPNRVGLPINQRFILVEL 62  
QY 54 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAGGN 113  
DB 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAGGN 122  
QY 114 YDRLEQLAGNREINELGNPLEEALSALYYSTGTQTLPTLARSFIIICQMISEAARFQ 173

DB 123 YDRLEQLAGNREINELGNPLEEALSALYYSTGTQTLPTLARSFIIICQMISEAARFQ 182  
QY 174 YIEGEMTRIRYNRS 189  
DB 183 YIEGEMTRIRYNRS 198  
RESULT 13  
US-10-083-336A-9  
Sequence 9, Application US/10083336A  
Publication No. US20030181665A1  
GENERAL INFORMATION:  
APPLICANT: Olson, Mark A  
APPLICANT: Millard, Charles B  
APPLICANT: Byrne, Michael P  
APPLICANT: Wannemacher, Robert W  
TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof  
FILE REFERENCE: P67452US0 (RIID 01-58)  
CURRENT APPLICATION NUMBER: US/10/083,336A  
CURRENT FILING DATE: 2002-05-21  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 185  
TYPE: PRT  
ORGANISM: Ricinus communis  
US-10-083-336A-9

Query Match 96.4%; Score 939.5; DB 12; Length 185;  
Best Local Similarity 98.4%; Pred. No. 1.2e-99;  
Matches 185; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTVLPNRVGLPINQRFILVELSNHAELSV 61  
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT--NRVGLPINQRFILVELSNHAELSV 57  
QY 62 TLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAGGNVDRLEOLA 121  
DB 58 TLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAGGNVDRLEOLA 117  
QY 122 GNREINELGNPLEEALSALYYSTGTQTLPTLARSFIIICQMISEAARFQYIEGEMRT 181  
DB 118 GNREINELGNPLEEALSALYYSTGTQTLPTLARSFIIICQMISEAARFQYIEGEMRT 177  
QY 182 RIRYNRS 189  
DB 178 RIRYNRS 185

RESULT 14  
US-10-083-336A-2  
Sequence 2, Application US/10083336A  
Publication No. US20030181665A1  
GENERAL INFORMATION:  
APPLICANT: Olson, Mark A  
APPLICANT: Millard, Charles B  
APPLICANT: Byrne, Michael P  
APPLICANT: Wannemacher, Robert W  
TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof  
FILE REFERENCE: P67452US0 (RIID 01-58)  
CURRENT APPLICATION NUMBER: US/10/083,336A  
CURRENT FILING DATE: 2002-05-21  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 179  
TYPE: PRT  
ORGANISM: Ricinus communis  
US-10-083-336A-2

Query Match 69.6%; Score 679; DB 12; Length 179;  
Best Local Similarity 93.1%; Pred. No. 9.9e-70;  
Matches 134; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

```
QY      2  IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLT-----VLPNRVGLPINQRFILV 51
Db      36  IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTGDVRHEIPVLPNRVGLPINQRFILV 95
QY     52  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQNOEDAETHLFTDVQNRITYFAFG 111
Db     96  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQNOEDAETHLFTDVQNRITYFAFG 155
QY    112  GNYDRLEQLAGNLRNIELNGNPL 135
Db    156  GNYDRLEQLAGNLRNIELNGNPL 179
```

## RESULT 15

```
US-09-792-793A-39
; Sequence 39, Application US/09792793A
; Patent No. US20020168370A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
; APPLICANT: Coggins, Philip
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND
; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
; FILE REFERENCE: 25020-601D
; CURRENT APPLICATION NUMBER: US/09/792,793A
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Trichosantheus kirilowii
US-09-792-793A-39
```

```
Query Match      34.5%; Score 336; DB 10; Length 247;
Best Local Similarity 38.9%; Pred. No. 3.9e-30;
Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;

QY      10  INFETAGATVQSYTNFIRAVGRGLTVLPN-----RVGLPINQRFILVELSNHAEL 59
Db      2  VSERLSGATSSSYGVFISNR---KALPNERKLYDIPLRSSLPQSQRVALIHLITNVADE 58
QY     60  SVTLALDVTNAYVVGVRAGNSAYFFHPDQNOEDA-EAITHLFTDVQNRITYFAFGGNYDRLE 118
Db     59  TISVALDVTNVIIMGVYRAGDTSYFF--NEASATEAAKVVKDAMRKVTLPYSGNYERLQ 115
QY    119  QLAGNLRNIELNGNPLEEASIALYVYSTGCTQLPLARSFIICIMISEAARFQYIEGE 178
Db    116  TAAGKIRENIPLGLPALDSAITTLFYNN-----SAASALMVLIQSTSEAAFYKFEIQ 170
QY    179  MRTRI 183
Db    171  IGRV 175
```

Search completed: February 10, 2004, 16:53:56  
Job time : 25.507 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:35 ; Search time 10.2817 Seconds  
(without alignments)  
1777.145 Million cell updates/sec

Title: US-10-083-336A-11

Perfect score: 975

Sequence: 1 MIFPKQYPIINFITAGATVQ.....RFQYIEGEMRTRINRNSA 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	955	97.9	576	1	RLCSD
2	865.5	88.8	564	1	RLCSAG
3	336	34.5	289	1	RLTGT
4	327.5	33.6	528	2	S32431
5	327.5	33.6	562	2	S16022
6	326	33.4	247	2	JU0393
7	326	33.4	247	2	JC5032
8	326	33.4	289	2	JC5006
9	323	33.1	527	2	S32430
10	310.5	31.8	251	2	C39761
11	303.5	31.1	528	1	TZLSA
12	300.5	30.8	278	2	S23519
13	293.5	30.1	250	2	JN0108
14	274.5	28.2	277	2	S22494
15	273	28.0	286	1	RLPUGG
16	273	28.0	570	2	S62627
17	272	27.9	286	2	S25560
18	270	27.7	245	2	JC4840
19	265	27.2	286	2	JC4235
20	264	27.1	254	2	PD0018
21	246.5	25.3	316	2	JT0753
22	186.5	19.1	294	2	S28421
23	174	17.8	278	2	A39817
24	170	17.4	313	2	S17757
25	168.5	17.3	261	2	JE0401
26	148.5	15.2	289	2	TJ2573
27	138	14.2	272	2	JC4811
28	131	13.4	253	2	S28542
29	127	13.0	253	2	S28539

## ALIGNMENTS

## RESULT 1

## RLCSD

ricin D precursor - castor bean

N;Contains: rRNA N-glycosidase (EC 3.2.2.22)

C;Species: Ricinus communis (castor bean)

C;Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 16-Jul-1999

C;Accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903

R;Halling, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.F.

Nucleic Acids Res. 13, 8019-8033, 1985

A;Title: Genomic cloning and characterization of a ricin gene from Ricinus communis.

A;Reference number: A24041; MUID:86067214; PMID:2399712

A;Accession: A24041

A;Molecule type: DNA

A;Residues: 1-576 <HAL>

A;Cross-references: GB:X03179; NID:g21082; PIDN:CAA26939.1; PID:g21083

R;Regear, J.W.; Roberts, L.M.

Plant Mol. Biol. 18, 515-525, 1992

A;Title: The lectin gene family of Ricinus communis: cloning of a functional ricin gene

A;Reference number: S20513; MUID:92163016; PMID:1371405

A;Accession: S20513

A;Molecule type: DNA

A;Residues: 1-576 <TRE>

A;Cross-references: EMBL:X52908; NID:g21084; PIDN:CAA37095.1; PID:g21085

R;Lamb, F.I.; Roberts, L.M.; Lord, J.M.

Eur. J. Biochem. 148, 265-270, 1985

A;Title: Nucleotide sequence of cloned cDNA coding for preprorcin.

A;Reference number: A24614; MUID:85179479; PMID:3838723

A;Accession: A24614

A;Molecule type: mRNA

A;Residues: 12-75, 'D', 77-550, 'R', 552-576 <LAM>

A;Cross-references: GB:X02388; NID:g21077; PIDN:CAA26230.1; PID:g21078

R;Yoshitake, S.; Funatsu, G.; Funatsu, M.

Agric. Biol. Chem. 42, 1267-1274, 1978

A;Title: Isolation and sequences of peptic peptides, and the complete sequence of Ile che

A;Reference number: A03372

A;Accession: A03372

A;Molecule type: protein

A;Residues: 36-97, 'Q', 99-109, 'S', 111-269, 'D', 272-283, 'L', 285-288, 290-302 <YOS>

A;Note: this paper cites the others in the series providing experimental details for the

R;Araki, T.; Funatsu, G.

FEBS Lett. 191, 121-124, 1985

A;Title: Revised amino acid sequence of the B-chain of ricin D due to loss of tryptophan

A;Reference number: A24010

A;Accession: A24010

A;Molecule type: protein

A;Residues: 315-383, 'PS', 386-576 <ARA>

R;Funatsu, G.; Kimura, M.; Funatsu, M.

Agric. Biol. Chem. 43, 2221-2224, 1979

A;Title: Primary structure of Ala chain of ricin D.

A;Reference number: A03374

A;Accession: A03374

A;Molecule type: protein

A:Residues: 315-335, 'N', 337-342, 'NH', 345-362, 364-383, 'PS', 386-399, 'T', 401, 'D', 403, 'E', 405-407, 'V', 409-410, 'L', 412-413, 'I', 415-416, 'F', 418-419, 'Y', 421-422, 'M', 424-425, 'K', 427-428, 'R', 430-431, 'G', 433-434, 'H', 436-437, 'Q', 439-440, 'E', 442-443, 'D', 445-446, 'N', 448-449, 'K', 451-452, 'R', 454-455, 'G', 457-458, 'H', 460-461, 'L', 463-464, 'I', 466-467, 'F', 469-470, 'Y', 472-473, 'M', 475-476, 'K', 478-479, 'R', 481-482, 'G', 484-485, 'H', 487-488, 'Q', 490-491, 'E', 493-494, 'D', 496-497, 'N', 499-500, 'K', 502-503, 'R', 505-506, 'G', 508-509, 'H', 511-512, 'L', 514-515, 'I', 517-518, 'F', 520-521, 'Y', 523-524, 'M', 526-527, 'K', 529-530, 'R', 532-533, 'G', 535-536, 'H', 538-539, 'Q', 541-542, 'E', 544-545, 'D', 547-548, 'N', 550-551, 'K', 553-554, 'R', 556-557, 'G', 559-560, 'H', 562-563, 'L', 565-566, 'I', 568-569, 'F', 570-571, 'Y', 573-574, 'M', 576-577, 'K', 579-580, 'R', 582-583, 'G', 585-586, 'H', 588-589, 'Q', 591-592, 'E', 594-595, 'D', 597-598, 'N', 600-601, 'K', 603-604, 'R', 606-607, 'G', 609-610, 'H', 612-613, 'L', 615-616, 'I', 618-619, 'F', 621-622, 'Y', 624-625, 'M', 627-628, 'K', 630-631, 'R', 633-634, 'G', 636-637, 'H', 639-640, 'Q', 642-643, 'E', 645-646, 'D', 648-649, 'N', 651-652, 'K', 654-655, 'R', 657-658, 'G', 660-661, 'H', 663-664, 'L', 666-667, 'I', 669-670, 'F', 672-673, 'Y', 675-676, 'M', 678-679, 'K', 681-682, 'R', 684-685, 'G', 687-688, 'H', 690-691, 'Q', 693-694, 'E', 696-697, 'D', 699-700, 'N', 702-703, 'K', 705-706, 'R', 708-709, 'G', 711-712, 'H', 714-715, 'L', 717-718, 'I', 720-721, 'F', 723-724, 'Y', 726-727, 'M', 729-730, 'K', 732-733, 'R', 735-736, 'G', 738-739, 'H', 741-742, 'Q', 744-745, 'E', 747-748, 'D', 750-751, 'N', 753-754, 'K', 756-757, 'R', 759-760, 'G', 762-763, 'H', 765-766, 'L', 768-769, 'I', 771-772, 'F', 774-775, 'Y', 777-778, 'M', 780-781, 'K', 783-784, 'R', 786-787, 'G', 789-790, 'H', 792-793, 'Q', 795-796, 'E', 798-799, 'D', 801-802, 'N', 804-805, 'K', 807-808, 'R', 810-811, 'G', 813-814, 'H', 816-817, 'L', 819-820, 'I', 822-823, 'F', 825-826, 'Y', 828-829, 'M', 831-832, 'K', 834-835, 'R', 837-838, 'G', 840-841, 'H', 843-844, 'Q', 846-847, 'E', 849-850, 'D', 852-853, 'N', 855-856, 'K', 858-859, 'R', 861-862, 'G', 864-865, 'H', 867-868, 'L', 869-870, 'I', 872-873, 'F', 875-876, 'Y', 878-879, 'M', 881-882, 'K', 883-884, 'R', 886-887, 'G', 889-890, 'H', 892-893, 'Q', 895-896, 'E', 898-899, 'D', 901-902, 'N', 904-905, 'K', 907-908, 'R', 910-911, 'G', 913-914, 'H', 916-917, 'L', 919-920, 'I', 922-923, 'F', 925-926, 'Y', 928-929, 'M', 931-932, 'K', 934-935, 'R', 937-938, 'G', 940-941, 'H', 943-944, 'Q', 946-947, 'E', 949-950, 'D', 952-953, 'N', 955-956, 'K', 958-959, 'R', 961-962, 'G', 964-965, 'H', 967-968, 'L', 969-970, 'I', 972-973, 'F', 975-976, 'Y', 978-979, 'M', 981-982, 'K', 983-984, 'R', 986-987, 'G', 989-990, 'H', 992-993, 'Q', 995-996, 'E', 998-999, 'D', 1001-1002, 'N', 1004-1005, 'K', 1007-1008, 'R', 1010-1011, 'G', 1013-1014, 'H', 1016-1017, 'L', 1019-1020, 'I', 1022-1023, 'F', 1025-1026, 'Y', 1028-1029, 'M', 1031-1032, 'K', 1034-1035, 'R', 1037-1038, 'G', 1040-1041, 'H', 1043-1044, 'Q', 1046-1047, 'E', 1049-1050, 'D', 1052-1053, 'N', 1055-1056, 'K', 1058-1059, 'R', 1061-1062, 'G', 1064-1065, 'H', 1067-1068, 'L', 1069-1070, 'I', 1072-1073, 'F', 1075-1076, 'Y', 1078-1079, 'M', 1081-1082, 'K', 1083-1084, 'R', 1086-1087, 'G', 1089-1090, 'H', 1092-1093, 'Q', 1095-1096, 'E', 1098-1099, 'D', 1101-1102, 'N', 1104-1105, 'K', 1107-1108, 'R', 1110-1111, 'G', 1113-1114, 'H', 1116-1117, 'L', 1119-1120, 'I', 1122-1123, 'F', 1125-1126, 'Y', 1128-1129, 'M', 1131-1132, 'K', 1134-1135, 'R', 1137-1138, 'G', 1140-1141, 'H', 1143-1144, 'Q', 1146-1147, 'E', 1149-1150, 'D', 1152-1153, 'N', 1155-1156, 'K', 1158-1159, 'R', 1161-1162, 'G', 1164-1165, 'H', 1167-1168, 'L', 1169-1170, 'I', 1172-1173, 'F', 1175-1176, 'Y', 1178-1179, 'M', 1181-1182, 'K', 1183-1184, 'R', 1186-1187, 'G', 1189-1190, 'H', 1192-1193, 'Q', 1195-1196, 'E', 1198-1199, 'D', 1201-1202, 'N', 1204-1205, 'K', 1207-1208, 'R', 1210-1211, 'G', 1213-1214, 'H', 1216-1217, 'L', 1219-1220, 'I', 1222-1223, 'F', 1225-1226, 'Y', 1228-1229, 'M', 1231-1232, 'K', 1234-1235, 'R', 1237-1238, 'G', 1240-1241, 'H', 1243-1244, 'Q', 1246-1247, 'E', 1249-1250, 'D', 1252-1253, 'N', 1255-1256, 'K', 1258-1259, 'R', 1261-1262, 'G', 1264-1265, 'H', 1267-1268, 'L', 1269-1270, 'I', 1272-1273, 'F', 1275-1276, 'Y', 1278-1279, 'M', 1281-1282, 'K', 1283-1284, 'R', 1286-1287, 'G', 1289-1290, 'H', 1292-1293, 'Q', 1295-1296, 'E', 1298-1299, 'D', 1301-1302, 'N', 1304-1305, 'K', 1307-1308, 'R', 1310-1311, 'G', 1313-1314, 'H', 1316-1317, 'L', 1319-1320, 'I', 1322-1323, 'F', 1325-1326, 'Y', 1328-1329, 'M', 1331-1332, 'K', 1334-1335, 'R', 1337-1338, 'G', 1340-1341, 'H', 1343-1344, 'Q', 1346-1347, 'E', 1349-1350, 'D', 1352-1353, 'N', 1355-1356, 'K', 1358-1359, 'R', 1361-1362, 'G', 1364-1365, 'H', 1367-1368, 'L', 1369-1370, 'I', 1372-1373, 'F', 1375-1376, 'Y', 1378-1379, 'M', 1381-1382, 'K', 1383-1384, 'R', 1386-1387, 'G', 1389-1390, 'H', 1392-1393, 'Q', 1395-1396, 'E', 1398-1399, 'D', 1401-1402, 'N', 1404-1405, 'K', 1407-1408, 'R', 1410-1411, 'G', 1413-1414, 'H', 1416-1417, 'L', 1419-1420, 'I', 1422-1423, 'F', 1425-1426, 'Y', 1428-1429, 'M', 1431-1432, 'K', 1434-1435, 'R', 1437-1438, 'G', 1440-1441, 'H', 1443-1444, 'Q', 1446-1447, 'E', 1449-1450, 'D', 1452-1453, 'N', 1455-1456, 'K', 1458-1459, 'R', 1461-1462, 'G', 1464-1465, 'H', 1467-1468, 'L', 1469-1470, 'I', 1472-1473, 'F', 1475-1476, 'Y', 1478-1479, 'M', 1481-1482, 'K', 1483-1484, 'R', 1486-1487, 'G', 1489-1490, 'H', 1492-1493, 'Q', 1495-1496, 'E', 1498-1499, 'D', 1501-1502, 'N', 1504-1505, 'K', 1507-1508, 'R', 1510-1511, 'G', 1513-1514, 'H', 1516-1517, 'L', 1519-1520, 'I', 1522-1523, 'F', 1525-1526, 'Y', 1528-1529, 'M', 1531-1532, 'K', 1534-1535, 'R', 1537-1538, 'G', 1540-1541, 'H', 1543-1544, 'Q', 1546-1547, 'E', 1549-1550, 'D', 1552-1553, 'N', 1555-1556, 'K', 1558-1559, 'R', 1561-1562, 'G', 1564-1565, 'H', 1567-1568, 'L', 1569-1570, 'I', 1572-1573, 'F', 1575-1576, 'Y', 1578-1579, 'M', 1581-1582, 'K', 1583-1584, 'R', 1586-1587, 'G', 1589-1590, 'H', 1592-1593, 'Q', 1595-1596, 'E', 1598-1599, 'D', 1601-1602, 'N', 1604-1605, 'K', 1607-1608, 'R', 1610-1611, 'G', 1613-1614, 'H', 1616-1617, 'L', 1619-1620, 'I', 1622-1623, 'F', 1625-1626, 'Y', 1628-1629, 'M', 1631-1632, 'K', 1634-1635, 'R', 1637-1638, 'G', 1640-1641, 'H', 1643-1644, 'Q', 1646-1647, 'E', 1649-1650, 'D', 1652-1653, 'N', 1655-1656, 'K', 1658-1659, 'R', 1661-1662, 'G', 1664-1665, 'H', 1667-1668, 'L', 1669-1670, 'I', 1672-1673, 'F', 1675-1676, 'Y', 1678-1679, 'M', 1681-1682, 'K', 1683-1684, 'R', 1686-1687, 'G', 1689-1690, 'H', 1692-1693, 'Q', 1695-1696, 'E', 1698-1699, 'D', 1701-1702, 'N', 1704-1705, 'K', 1707-1708, 'R', 1710-1711, 'G', 1713-1714, 'H', 1716-1717, 'L', 1719-1720, 'I', 1722-1723, 'F', 1725-1726, 'Y', 1728-1729, 'M', 1731-1732, 'K', 1734-1735, 'R', 1737-1738, 'G', 1740-1741, 'H', 1743-1744, 'Q', 1746-1747, 'E', 1749-1750, 'D', 1752-1753, 'N', 1755-1756, 'K', 1758-1759, 'R', 1761-1762, 'G', 1764-1765, 'H', 1767-1768, 'L', 1769-1770, 'I', 1772-1773, 'F', 1775-1776, 'Y', 1778-1779, 'M', 1781-1782, 'K', 1783-1784, 'R', 1786-1787, 'G', 1789-1790, 'H', 1792-1793, 'Q', 1795-1796, 'E', 1798-1799, 'D', 1801-1802, 'N', 1804-1805, 'K', 1807-1808, 'R', 1810-1811, 'G', 1813-1814, 'H', 1816-1817, 'L', 1819-1820, 'I', 1822-1823, 'F', 1825-1826, 'Y', 1828-1829, 'M', 1831-1832, 'K', 1834-1835, 'R', 1837-1838, 'G', 1840-1841, 'H', 1843-1844, 'Q', 1846-1847, 'E', 1849-1850, 'D', 1852-1853, 'N', 1855-1856, 'K', 1858-1859, 'R', 1861-1862, 'G', 1864-1865, 'H', 1867-1868, 'L', 1869-1870, 'I', 1872-1873, 'F', 1875-1876, 'Y', 1878-1879, 'M', 1881-1882, 'K', 1883-1884, 'R', 1886-1887, 'G', 1889-1890, 'H', 1892-1893, 'Q', 1895-1896, 'E', 1898-1899, 'D', 1901-1902, 'N', 1904-1905, 'K', 1907-1908, 'R', 1910-1911, 'G', 1913-1914, 'H', 1916-1917, 'L', 1919-1920, 'I', 1922-1923, 'F', 1925-1926, 'Y', 1928-1929, 'M', 1931-1932, 'K', 1934-1935, 'R', 1937-1938, 'G', 1940-1941, 'H', 1943-1944, 'Q', 1946-1947, 'E', 1949-1950, 'D', 1952-1953, 'N', 1955-1956, 'K', 1958-1959, 'R', 1961-1962, 'G', 1964-1965, 'H', 1967-1968, 'L', 1969-1970, 'I', 1972-1973, 'F', 1975-1976, 'Y', 1978-1979, 'M', 1981-1982, 'K', 1983-1984, 'R', 1986-1987, 'G', 1989-1990, 'H', 1992-1993, 'Q', 1995-1996, 'E', 1998-1999, 'D', 2001-2002, 'N', 2004-2005, 'K', 2007-2008, 'R', 2010-2011, 'G', 2013-2014, 'H', 2016-2017, 'L', 2019-2020, 'I', 2022-2023, 'F', 2025-2026, 'Y', 2028-2029, 'M', 2031-2032, 'K', 2034-2035, 'R', 2037-2038, 'G', 2040-2041, 'H', 2043-2044, 'Q', 2046-2047, 'E', 2049-2050, 'D', 2052-2053, 'N', 2055-2056, 'K', 2058-2059, 'R', 2061-2062, 'G', 2064-2065, 'H', 2067-2068, 'L', 2069-2070, 'I', 2072-2073, 'F', 2075-2076, 'Y', 2078-2079, 'M', 2081-2082, 'K', 2083-2084, 'R', 2086-2087, 'G', 2089-2090, 'H', 2092-2093, 'Q', 2095-2096, 'E', 2098-2099, 'D', 2101-2102, 'N', 2104-2105, 'K', 2107-2108, 'R', 2110-2111, 'G', 2113-2114, 'H', 2116-2117, 'L', 2119-2120, 'I', 2122-2123, 'F', 2125-2126, 'Y', 2128-2129, 'M', 2131-2132, 'K', 2134-2135, 'R', 2137-2138, 'G', 2140-2141, 'H', 2143-2144, 'Q', 2146-2147, 'E', 2149-2150, 'D', 2152-2153, 'N', 2155-2156, 'K', 2158-2159, 'R', 2161-2162, 'G', 2164-2165, 'H', 2167-2168, 'L', 2169-2170, 'I', 2172-2173, 'F', 2175-2176, 'Y', 2178-2179, 'M', 2181-2182, 'K', 2183-2184, 'R', 2186-2187, 'G', 2189-2190, 'H', 2192-2193, 'Q', 2195-2196, 'E', 2198-2199, 'D', 2201-2202, 'N', 2204-2205, 'K', 2207-2208, 'R', 2210-2211, 'G', 2213-2214, 'H', 2216-2217, 'L', 2219-2220, 'I', 2222-2223, 'F', 2225-2226, 'Y', 2228-2229, 'M', 2231-2232, 'K', 2234-2235, 'R', 2237-2238, 'G', 2240-2241, 'H', 2243-2244, 'Q', 2246-2247, 'E', 2249-2250, 'D', 2252-2253, 'N', 2255-2256, 'K', 2258-2259, 'R', 2261-2262, 'G', 2264-2265, 'H', 2267-2268, 'L', 2269-2270, 'I', 2272-2273, 'F', 2275-2276, 'Y', 2278-2279, 'M', 2281-2282, 'K', 2283-2284, 'R', 2286-2287, 'G', 2289-2290, 'H', 2292-2293, 'Q', 2295-2296, 'E', 2298-2299, 'D', 2301-2302, 'N', 2304-2305, 'K', 2307-2308, 'R', 2310-2311, 'G', 2313-2314, 'H', 2316-2317, 'L', 2319-2320, 'I', 2322-2323, 'F', 2325-2326, 'Y', 2328-2329, 'M', 2331-2332, 'K', 2334-2335, 'R', 2337-2338, 'G', 2340-2341, 'H', 2343-2344, 'Q', 2346-2347, 'E', 2349-2350, 'D', 2352-2353, 'N', 2355-2356, 'K', 2358-2359, 'R', 2361-2362, 'G', 2364-2365, 'H', 2367-2368, 'L', 2369-2370, 'I', 2372-2373, 'F', 2375-2376, 'Y', 2378-2379, 'M', 2381-2382, 'K', 2383-2384, 'R', 2386-2387, 'G', 2389-2390, 'H', 2392-2393, 'Q', 2395-2396, 'E', 2398-2399, 'D', 2401-2402, 'N', 2404-2405, 'K', 2407-2408, 'R', 2410-2411, 'G', 2413-2414, 'H', 2416-2417, 'L', 2419-2420, 'I', 2422-2423, 'F', 2425-2426, 'Y', 2428-2429, 'M', 2431-2432, 'K', 2434-2435, 'R', 2437-2438, 'G', 2440-2441, 'H', 2443-2444, 'Q', 2446-2447, 'E', 2449-2450, 'D', 2452-2453, 'N', 2455-2456, 'K', 2458-2459, 'R', 2461-2462, 'G', 2464-2465, 'H', 2467-2468, 'L', 2469-2470, 'I', 2472-2473, 'F', 2475-2476, 'Y', 2478-2479, 'M', 2481-2482, 'K', 2483-2484, 'R', 2486-2487, 'G', 2489-2490, 'H', 2492-2493, 'Q', 2495-2496, 'E', 2498-2499, 'D', 2501-2502, 'N', 2504-2505, 'K', 2507-2508, 'R', 2510-2511, 'G', 2513-2514, 'H', 2516-2517, 'L', 2519-2520, 'I', 2522-2523, 'F', 2525-2526, 'Y', 2528-2529, 'M', 2531-2532, 'K', 2534-2535, 'R', 2537-2538, 'G', 2540-2541, 'H', 2543-2544, 'Q', 2546-2547, 'E', 2549-2550, 'D', 2552-2553, 'N', 2555-2556, 'K', 2558-2559, 'R', 2561-2562, 'G', 2564-2565, 'H', 2567-2568, 'L', 2569-2570, 'I', 2572-2573, 'F', 2575-2576, 'Y', 2578-2579, 'M', 2581-2582, 'K', 2583-2584, 'R', 2586-2587, 'G', 2589-2590, 'H', 2592-2593, 'Q', 2595-2596, 'E', 2598-2599, 'D', 2601-2602, 'N', 2604-2605, 'K', 2607-2608, 'R', 2610-2611, 'G', 2613-2614, 'H', 2616-2617, 'L', 2619-2620, 'I', 2622-2623, 'F', 2625-2626, 'Y', 2628-2629, 'M', 2631-2632, 'K', 2634-2635, 'R', 2637-2638, 'G', 2640-2641, 'H', 2643-2644, 'Q', 2646-2647, 'E', 2649-2650, 'D', 2652-2653, 'N', 2655-2656, 'K', 2658-2659, 'R', 2661-2662, 'G', 2664-2665, 'H', 2667-2668, 'L', 2669-2670, 'I', 2672-2673, 'F', 2675-2676, 'Y', 2678-2679, 'M', 2681-2682, 'K', 2683-2684, 'R', 2686-2687, 'G', 2689-2690, 'H', 2692-2693, 'Q', 2695-2696, 'E', 2698-2699, 'D', 2701-2702, 'N', 2704-2705, 'K', 2707-2708, 'R', 2710-2711, 'G', 2713-2714, 'H', 2716-2717, 'L', 2719-2720, 'I', 2722-2723, 'F', 2725-2726, 'Y', 2728-2729, 'M', 2731-2732, 'K', 2734-2735, 'R', 2737-2738, 'G', 2740-2741, 'H', 2743-2744, 'Q', 2746-2747, 'E', 2749-2750, 'D', 2752-2753, 'N', 2755-2756, 'K', 2758-2759, 'R', 2761-2762, 'G', 2764-2765, 'H', 2767-2768, 'L', 2769-2770, 'I', 2772-2773, 'F', 2775-2776, 'Y', 2778-2779, 'M', 2781-2782, 'K', 2783-2784, 'R', 2786-2787, 'G', 2789-2790, 'H', 2792-2793, 'Q', 2795-2796, 'E', 2798-2799, 'D', 2801-2802, 'N', 2804-2805, 'K', 2807-2808, 'R', 2810-2811, 'G', 2813-2814, 'H', 2816-2817, 'L', 2819-2820, 'I', 2822-2823, 'F', 2825-2826, 'Y', 2828-2829, 'M', 2831-2832, 'K', 2834-2835, 'R', 2837-2838, 'G', 2840-2841, 'H', 2843-2844, 'Q', 2846-2847, 'E', 2849-2850, 'D', 2852-2853, 'N', 2855-2856, 'K', 2858-2859, 'R', 2861-2862, 'G', 2864-2865, 'H', 2867-2868, 'L', 2869-2870, 'I', 2872-2873, 'F', 2875-2876, 'Y', 2878-2879, 'M', 2881-2882, 'K', 2883-2884, 'R', 2886-2887, 'G', 2889-2890, 'H', 2892-2893, 'Q', 2895-2896, 'E', 2898-2899, 'D', 2901-2902, 'N', 2904-2905, 'K', 2907-2908, 'R', 2910-2911, 'G', 2913-2914, 'H', 2916-2917, 'L', 2919-2920, 'I', 2922-2923, 'F', 2925-2926, 'Y', 2928-2929, 'M', 2931-2932, 'K', 2934-2935, 'R', 2937-2938, 'G', 2940-2941, 'H', 2943-2944, 'Q', 2946-2947, 'E', 2949-2950, 'D', 2952-2953, 'N', 2955-2956, 'K', 2958-2959, 'R', 2961-2962, 'G', 2964-2965, 'H', 2967-2968, 'L', 2969-2970, 'I', 2972-2973, 'F', 2975-2976, 'Y', 2978-2979, 'M', 2981-2982, 'K', 2983-2984, 'R', 2986-2987, 'G', 2989-2990, 'H', 2992-2993, 'Q', 2995-2996, 'E', 2998-2999, 'D', 3001-3002, 'N', 3004-3005, 'K', 3007-3008, 'R', 3010-3011, 'G', 3013-3014, 'H', 3016-3017, 'L', 3019-3020, 'I', 3022-3023, 'F', 3025-3026, 'Y', 3028-3029, 'M', 3031-3032, 'K', 3034-3035, 'R', 3037-3038, 'G', 3040-3041, 'H', 3043-3044, 'Q', 3046-3047, 'E', 304



A:Molecule type: DNA

A:Residues: 1-72, 'V', 74-90, 'S', 92-233, 'T', 235-267, 'D', 269-289 <ZHE>  
A:Cross-references: GB:S70176; NID:G547148; PIDN:AAB31048.1; PID:G547149  
R:Collins, E.J.; Robertus, J.D.; Lopresti, M.; Stone, K.L.; Williams, K.R.; Wu, P.; Hwang, J. Biol. Chem. 265, 8665-8669, 1990

A:Title: Primary amino acid sequence of alpha-trichosanthin and molecular models for abrin

A:Reference number: A36273; MUID:90256789; PMID:2341399

A:Accession: A36273

A:Molecule type: protein

A:Residues: 24-270 <COL>

R:Wang, Y.; Qian, R.Q.; Gu, Z.W.; Jin, S.W.; Zhang, L.Q.; Xia, Z.X.; Tian, G.Y.; Ni, C.Z. Pure Appl. Chem. 58, 789-798, 1986

A:Title: Scientific evaluation of Tian Hua Fen (THF): history, chemistry and application

A:Reference number: JT0003

A:Accession: JT0003

A:Molecule type: protein

A:Residues: 24-56, 'L', 58-59, 'I', 61-71, 'I', 73-81, 85-86, 'L', 88-92, 'DAGLPNVAVL', 93-142, 'GL'

A:Experimental source: tuber

R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.

A:Submitted to the Brookhaven Protein Data Bank, July 1994

A:Reference number: A67091; PDB:IMR7

A:Contents: annotation; X-ray crystallography, 1.6 angstroms, with adenine, residues 24-

R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.

A:Submitted to the Brookhaven Protein Data Bank, July 1994

A:Reference number: A67092; PDB:IMR8

A:Contents: annotation; X-ray crystallography, 1.6 angstroms, with formycin, residues 24-

R:Xiong, J.P.; Xia, Z.X.; Wang, Y.

A:Submitted to the Brookhaven Protein Data Bank, December 1994

A:Reference number: A67111; PDB:ITCS

A:Contents: annotation; X-ray crystallography, 1.7 angstroms, with NADPH, residues 24-27

R:Xiong, J.P.; Xia, Z.X.; Wang, Y.

A:Submitted to the Brookhaven Protein Data Bank, July 1994

A:Title: Crystal structure of trichosanthin-NADPH complex at 1.7 Angstroms resolution re

A:Reference number: A58622; MUID:95360714; PMID:7634073

A:Contents: annotation; X-ray crystallography, 1.7 angstroms

A:Comment: Alpha-trichosanthin has been used to induce abortions.

C:Genetics:

A:Gene: tcs

C:Function:

A:Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA then

C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology

C:Keywords: abortifacient; glycosidase; hydrolase; root; toxin

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-270/Product: trichosanthin alpha #status experimental <MAT>

F:27-266/Domain: rRNA N-glycosidase homology <RNG>

F:271-289/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F:93,183,186/Active site: Tyr, Glu, Arg #status predicted

Query Match 34.5%; Score 336; DB 1; Length 289;

Best Local Similarity 38.9%; Pred. No. 2,8e-23;

Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;

QY 10 INFATTAGATVQSTNFIHRAVGRGLTVPN-----RVGLPINQRFILVELSNHAE 59

Db 25 VSPRLSGATSSSYGVFISNLR---KALPNKLYDIPLLRSSLPQSRYVALIHLTNVADE 81

QY 60 SVTLALDVTNAYVGVYGRAGNSAFPHDPNOEDA-EAITHLFTDVQNRVTFAGCNVDRL 118

Db 82 TTSVALDVTNAYVGVYGRAGNSAFPHDPNOEDA-EAITHLFTDVQNRVTFAGCNVDRL 138

QY 119 QLAGNLRENIENGLNGPLEEASALYTYSTGTQTLPLARSFFIICQWISSEARFQYIEG 178

Db 139 TAAGKIRENIPGLPALDSAITLTFYNN-----SAASALMVLIOSTSEARYKIEQ 193

QY 179 MRTRI 183

Db 194 IGRV 198

RESULT 4

S32431

abrin-d precursor - Indian licorice (fragment)

N:Contains: rRNA N-glycosidase (EC 3.2.2.22)

C:Species: Abrus precatorius (Indian licorice)

C:Date: 30-Sep-1993 #sequence\_revision 01-Aug-1997 #text\_change 01-Aug-1997

C:Accession: S32431; S34408

R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.

J. Mol. Biol. 229, 263-267, 1993

A:Title: Primary structure of three distinct isoabrin determined by cDNA sequencing. Cor

A:Reference number: S32429; MUID:93132798; PMID:8421131

A:Accession: S32431

A:Molecule type: mRNA

A:Residues: 1-528 <HUN>

A:Cross-references: GB:M98346

R:Hung, C.; Lee, M.; Lee, T.; Lin, J.

submitted to the EMBL Data Library, March 1993

A:Reference number: S34408

A:Accession: S34408

A:Molecule type: mRNA

A:Residues: 1-169, 'C', 171-320, 'L', 322-528 <HU2>

A:Cross-references: GB:M98346

C:Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating

The A and B chains are linked by a single disulfide bond, which is essential for toxicit

C:Superfamily: ricin; rRNA N-glycosidase homology

C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; p

F:1-251/Product: abrin-d chain A #status predicted <ACH>

F:7-245/Domain: rRNA N-glycosidase homology <RNG>

F:261-528/Product: abrin-d chain B #status predicted <BCH>

F:283-325,326-366,369-407,414-449,453-492,495-528/Region: 40-residue repeats

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted

F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted

F:164,167/Active site: Glu, Arg #status predicted

F:200,253,361,401,402/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:247-269,286-305,329-346,417-430,456-473/Disulfide bonds: #status predicted

F:288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

F:500,521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 33.6%; Score 327.5; DB 2; Length 528;

Best Local Similarity 43.9%; Pred. No. 3.6e-22;

Matches 82; Conservative 22; Mismatches 68; Indels 15; Gaps 4;

QY 6 QVPINFTTAGATVQSTNFIHRAVGRGLT-----VLPRVGLPINQRFILVELSNHA 57

Db 1 QQVVKFTTEGATSSQYKQFIEALQRLTGGLIHDPVLPDPTTVEERNRYITVELSNSE 60

QY 58 ELSVTALDVTNAYVGVYGRAGNSAFPHDPNOEDAEATHLFTDVQNRVTFAGCNVDRL 117

Db 61 RESIEVGIDVTNAYVGVYGRAGNSAFPHDPNOEDAEATHLFTDVQNRVTFAGCNVDRL 116

QY 118 EQLAGNLRENIENGLNGPLEEASALYTYSTGTQTLPLARSFFIICQWISSEARFQYIEG 177

Db 117 ERWAHQTRREISLGLQALTHAIS---FLRSGASNDDEKARTLIVIQMASEARVRYISN 173

QY 178 EMRTIR 184

Db 174 RVGVSR 180

RESULT 5

S16022

abrin-c precursor - Indian licorice

N:Contains: rRNA N-glycosidase (EC 3.2.2.22)

C:Species: Abrus precatorius (Indian licorice)

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Aug-1999

C:Accession: S16022

R:Wood, K.A.; Lord, J.M.; Wawrzynczak, E.J.; Piatak, M.

Eur. J. Biochem. 198, 723-732, 1991

A:Title: Preproabrin: genomic cloning, characterisation and the expression of the A-chain

A:Reference number: S16022; MUID:91266957; PMID:2050149

A:Accession: S16022

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-562 <WOO>

A:Cross-references: EMBL:X55667; NID:g16084; PIDN:CAA39202.1; PID:g16085

C:Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating

The A and B chains are linked by a single disulfide bond, which is essential for toxicit

C:Superfamily: ricin; rRNA N-glycosidase homology  
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglutamic acid  
F:35-285/Product: abrin-c chain A #status predicted <ACH>  
F:41-280/Domain: rRNA N-glycosidase homology <RNG>  
F:295-562/Product: abrin-c chain B #status predicted <BOH>  
F:317-359,360-400,403-441,448-483,487-526,529-562/Region: 40-residue repeats  
F:335/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:108,147,229,230/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
F:198,201/Active site: Glu, Arg #status predicted  
F:281,303,320-339,363-380,451-464,490-507/Disulfide bonds: #status predicted  
F:322,346/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted  
F:534,555/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

Query Match 33.6%; Score 327.5; DB 2; Length 562;  
Best Local Similarity 43.9%; Pred. No. 3.9e-22;  
Matches 82; Conservative 22; Mismatches 68; Indels 15; Gaps 4;

QY 6 QYPIINFTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILVELSNHA 57  
Db 35 QDQIKFTFEGATQSYKQFIALRQRLTGGLIHDPDPTTVEERNRYITVELSNSE 94  
QY 58 ELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFGNRYDL 117  
Db 95 RESIEVGIDVTNAYVVRAGNSQSYFL---RDAPASASTYLPFGTQ-RYSLRFDSYGD 150  
QY 118 EQLAGNLENELNGPLEEALSALYYSTGTQPTLARSFIIQIMISEARFOYIEG 177  
Db 151 ERWQRTREISLGLQALTHAIS---FLRSGASNDDEKARTLIVIQASEAARYIYSN 207  
QY 178 EMRTIR 184  
Db 208 RVGSIR 214

RESULT 6  
JU0393  
karasurin - Mongolian snake-gourd  
C:Species: Trichosanthes kirilowii (Mongolian snake-gourd)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 12-Apr-1995  
C:Accession: JU0393; PS0163  
R:Toyokawa, S.; Takeda, T.; Kato, Y.; Wakabayashi, K.; Ogiwara, Y.  
Chem. Pharm. Bull. 39, 1244-1249, 1991  
A:Title: The complete amino acid sequence of an abortifacient protein, karasurin.  
A:Reference number: JU0393; MUID:92005321; PMID:1914000  
A:Accession: JU0393  
A:Molecule type: protein  
A:Residues: 1-247 <TOY>  
A:Note: a sequence which lacks Ala-247 is also shown in this publication  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C:Keywords: abortifacient  
F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 33.4%; Score 326; DB 2; Length 247;  
Best Local Similarity 39.6%; Pred. No. 1.9e-22;  
Matches 72; Conservative 44; Mismatches 50; Indels 16; Gaps 5;

QY 10 INFTTAGATVQSYTNFIRAVR-----GRLTVLP-NRVGLPINQRFILVELSNHAEISVT 62  
Db 2 VSRFLSGATSSSYGVFISNLKALPYERKLYDIPLLRSTLPGSQRYALIHLTNYADETIS 61  
QY 63 LALDVTNAYVVGVRAGNSAYFFHPDNOEDA-EATHLFTDVQNRYYTFAFGNRYDL 121  
Db 62 VAIDVTNVVVGVRAGDTSYFF---NEASATEAAKYVFKDKRKVTLPYSGNYERLQIAA 118  
QY 122 GNLENELNGPLEEALSALYYSTGTQPTLARSFIIQIMISEARFOYIEGEMRT 181  
Db 119 GKIRENIPGLPALDSAITTLFFYNAN-----SAASALMWLIQSTSEARYKFIQQIGK 173  
QY 182 RI 183  
Db 174 RV 175

RESULT 7  
JC5032  
karasurin-B - Trichosanthes kirilowii var. japonica  
C:Species: Trichosanthes kirilowii var. japonica  
C:Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 23-May-1997  
C:Accession: JC5032  
R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogiwara, Y.  
Biolog. Pharm. Bull. 19, 1485-1489, 1996  
A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and ka:  
A:Reference number: JC5032; MUID:97108848; PMID:8951169  
A:Accession: JC5032  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-247 <KON>  
A:Comment: This protein belongs to type I ribosomal-inactivating proteins which catalyti:  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 33.4%; Score 326; DB 2; Length 247;  
Best Local Similarity 39.6%; Pred. No. 1.9e-22;  
Matches 72; Conservative 44; Mismatches 50; Indels 16; Gaps 5;

QY 10 INFTTAGATVQSYTNFIRAVR-----GRLTVLP-NRVGLPINQRFILVELSNHAEISVT 62  
Db 2 VSRFLSGATSSSYGVFISNLKALPYERKLYDIPLLRSTLPGSQRYALIHLTNYADETIS 61  
QY 63 LALDVTNAYVVGVRAGNSAYFFHPDNOEDA-EATHLFTDVQNRYYTFAFGNRYDL 121  
Db 62 VAIDVTNVVVGVRAGDTSYFF---NEASATEAAKYVFKDKRKVTLPYSGNYERLQIAA 118  
QY 122 GNLENELNGPLEEALSALYYSTGTQPTLARSFIIQIMISEARFOYIEGEMRT 181  
Db 119 GKIRENIPGLPALDSAITTLFFYNAN-----SAASALMWLIQSTSEARYKFIQQIGK 173  
QY 182 RI 183  
Db 174 RV 175

RESULT 8  
JC5606  
karasurin C - Trichosanthes kirilowii var. japonica  
N:Contains: karasurin A  
C:Species: Trichosanthes kirilowii var. japonica  
C:Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 19-Jul-2002  
C:Accession: JC5606; JC5033  
R:Mizukami, H.; Iida, K.; Kondo, T.; Ogiwara, Y.  
Biolog. Pharm. Bull. 20, 711-713, 1997  
A:Title: Cloning and bacterial expression of a gene encoding ribosome-inactivating prote:  
A:Reference number: JC5606; MUID:97356562; PMID:9212998  
A:Accession: JC5606  
A:Molecule type: DNA  
A:Residues: 1-289 <Miz>  
A:Cross-references: DDBJ:AB000666; NID:G2329830; PIDN:BAA21786.1; PID:G2329831  
R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogiwara, Y.  
Biolog. Pharm. Bull. 19, 1485-1489, 1996  
A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and ka:  
A:Reference number: JC5032; MUID:97108848; PMID:8951169  
A:Accession: JC5033  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 22-270 <KON>  
A:Comment: This protein is a ribosome-inactivating protein and exhibit cytotoxic, aborti:  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
F:22-270/Product: karasurin C #status predicted <MAA>  
F:24-270/Product: karasurin A #status predicted <MAA>  
F:27-266/Domain: rRNA N-glycosidase homology <RNG>

Query Match 33.4%; Score 326; DB 2; Length 289;  
Best Local Similarity 39.6%; Pred. No. 2.3e-22;  
Matches 72; Conservative 44; Mismatches 50; Indels 16; Gaps 5;

```

QY 10 INFATTAGATVQSYNTNFAVR-----GRLTVLP-NRVGLPINQRFILVELSNHAEISVT 62
Db 25 VSLRSGATSSYGVFTSNLRKALPYERKLYDIPLRSTRFLGSGRYALIHUHTVADETIS 84
QY 63 LALDVNAVYVGYRAGNSAYFFHPDQDEA-EAITHLFTDVQNRVYTFAGNGYDRLEQLA 121
Db 85 VALDVNAVYVGYRAGDTSYFF---NEASATEAAKYVFKDAKKVTLPSYNGYERLQATA 141
QY 122 GNLRNRELNGPLUEEAISALYYSTGTGTQPLTARSFIICQMISEAARFQYIEGMRT 181
Db 142 GKIRENIPGLPALDSAITTLFFYNAN-----SAASALMWLIQSTSEAAKYKFFIEQIGK 196
QY 182 RI 183
Db 197 RV 198

RESULT 9
abrin-b precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C>Date: 30-Sep-1993 #sequence_revision 01-Aug-1997 #text_change 20-Aug-1999
C:Accession: S32430; JCI399
R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J. Mol. Biol. 229, 263-267, 1993
A:Title: Primary structure of three distinct isoabrin determined by cDNA sequencing.
A:Reference number: S32429; MUID:93132798; PMID:8421313
A:Accession: S32430
A:Molecule type: mRNA
A:Residues: 1-527 <HUN>
A:Cross-references: GB:M98345; NID:gl66296; PIDN:AAA32625.1; PID:gl66297.
R:Kimura, M.; Sumizawa, T.; Funatsu, G.
BioSci. Biotechnol. Biochem. 57, 166-169, 1993
A:Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxic
A:Reference number: JCI398; MUID:93169023; PMID:7763422
A:Accession: JCI399
A:Molecule type: Protein
A:Residues: 260-281, 'D', 283-290, 'N', 292-349, 'PQ', 352-377, 'N', 379-425, 'M', 427, 'D', 429-430
A:Experimental source: Seed
C:Superfamily: ricin; rRNA N-glycosidase homology
C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; P
F:1-250/Product: abrin-b chain A #status predicted <ACH>
F:7-245/Domain: rRNA N-glycosidase homology <RNG>
F:260-527/Product: abrin-b chain B #status experimental <BCH>
F:282-324,325-365,368-406,413-448,452-491,494-527/Region: 40-residue repeats
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted
F:74,113,194,195/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:110,360,400/Binding site: carboxylate (Asn) (covalent) #status predicted
F:163,166/Active site: Glu, Arg #status predicted
F:246-268,285-304,328-345,416-429,455-472/Disulfide bonds: #status predicted
F:287,311/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted
F:499,520/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 33.1%; Score 323; DB 2; Length 527;
Best Local Similarity 43.9%; Pred. No. 9, 2e-22;
Matches 83; Conservative 21; Mismatches 69; Indels 16; Gaps 4;

QY 6 QYPINFTTAGATVQSYNTNFAVRGRLT-----VLPNRVGLPINQRFILVELSNHA 57
Db 1 QDOVIFKFTTEGATSQSYKQFTEALRQLRGLTGLHIGIPVLPDPTTLQERNYISVELNSD 60
QY 58 ELSVTLADVNAVYVGYRAGNSAYFFHPDQDEAETHLFTDVQNRVYTFAGNGYDR 117
Db 61 TESLEAGIDVNAVYVGYRAGNSYFL-----RDAPTSASRYLFTGTQ-QYSLRFGSYIDL 116
QY 118 POLAGNLRNRELNGPLUEEAISALYYSTGTGTQPLTARSFIICQMISEAARFQYIEG 177
Db 117 ERLARQTFQQLPLGLQALRHAISEF-----QSGTDDQETARTLVIQWASEAARFRLSY 172
QY 178 EMRTIRYN 186

```

```

Db 173 RVGVSIRTN 181

RESULT 10
abrin (clone 7.2) precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C>Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 20-Aug-1999
C:Accession: C39761; S14471
R:Evensen, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266, 6848-6852, 1991
A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.
A:Reference number: A39761; MUID:91201329; PMID:2016300
A:Accession: C39761
A:Molecule type: DNA
A:Residues: 1-251 <EVE>
R:Evensen, G.; Mathiesen, A.; Sundan, A.
submitted to the EMBL Data Library, October 1990
A:Description: Direct molecular cloning of two distinct abrin A-chains.
A:Reference number: S14471
A:Accession: S14471
A:Molecule type: DNA
A:Residues: 'M', 1-251 <EV2>
A:Cross-references: EMBL:X54872; NID:gl6088; PIDN:CAA38654.1; PID:gl6089
C:Superfamily: ricin; rRNA N-glycosidase homology
C:Keywords: duplication; glycosidase; hydrolase; lectin; toxin
F:1-251/Product: abrin (clone 7.2) chain A #status predicted <ACH>
F:7-246/Domain: rRNA N-glycosidase homology <RNG>
F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:164,167/Active site: Glu, Arg #status predicted

Query Match 31.8%; Score 310.5; DB 2; Length 251;
Best Local Similarity 43.2%; Pred. No. 5e-21;
Matches 79; Conservative 22; Mismatches 67; Indels 15; Gaps 4;

QY 10 INFATTAGATVQSYNTNFAVRGRLT-----TVLPNRVGLPINQRFILVELSNHAEISV 61
Db 5 IKFTEGATSQSYKQFTEALRQLRGLIHIDIPVLRDPTTVEERNRYITVELSNSRESI 64
QY 62 TLALDVNAVYVGYRAGNSAYFFHPDQDEAETHLFTDVQNRVYTFAGNGYDRLEQLA 121
Db 65 EVGIDVNAVYVGYRAGNSQSYFL---RDAPASSTYLTGTQ-RYSRFDGSGYGLERWA 120
QY 122 GNLRNRELNGPLUEEAISALYYSTGTGTQPLTARSFIICQMISEAARFQYIEGMRT 181
Db 121 HQTREQISLGLQALTHAIS---FLRSGASNDKARTLVIQWASEAARFYISNRGV 177
QY 182 RIR 184
Db 178 SIR 180

RESULT 11
TZLSA
abrin-a precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C>Date: 31-Dec-1993 #sequence_revision 01-Aug-1997 #text_change 16-Jul-1999
C:Accession: S32429; JCI398; JCI398; S14472; S24133; S74110; S74111
R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J. Mol. Biol. 229, 263-267, 1993
A:Title: Primary structure of three distinct isoabrin determined by cDNA sequencing.
A:Reference number: S32429; MUID:93132798; PMID:8421313
A:Accession: S32429
A:Molecule type: mRNA
A:Residues: 'E', 2-528 <HUN>
A:Cross-references: GB:M98344; NID:gl66294; PIDN:AAA32624.1; PID:gl66295.
A:Note: the coding region for the sequence shown is preceded by an ATG codon
R:Funatsu, G.; Taguchi, Y.; Kamenosono, M.; Yanaka, M.
Agric. Biol. Chem. 52, 1095-1097, 1988

```

QY 10 INFTTAGATVQSYNNFTRAVRGRL-----TVLPNRVGLPINQRFILVELSNHAEISV 61  
 DB 5 IKFSTEGATSOYKQFTEALRERLGGIHDIPVLDPPTTLOERNRYITVELSNSDTSI 64  
 QY 62 TLALDVTNAYVVGVRAGNSAVFFH--PDNOEDAEATHLFTDVQNRVTFAPGGNYDRLEQ 119  
 DB 65 EVGIDVTNAYVVRAGTQSYFLRNDAPSSASD-----YLFGTGTDHSLDPFYGTYGDLR 118  
 QY 120 LAGNLRENIELGNGPLBEAISALYYVYTGGTQGLPTLARSPFIICIMISEAARFYIEGEM 179  
 DB 119 WAHQSRQOIPGLQALTHGIS---FFRSGGNDNEEKARTLIVIQWVAEAAARFYISNRV 175  
 QY 180 RTRIR 184  
 DB 176 RVSIQ 180  
  
 RESULT 12  
 S23519  
 beta-luffin - smooth loofah  
 C:Species: Luffa cylindrica (smooth loofah)  
 C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 20-Aug-1999  
 C:Accession: S23519; S23113  
 R:Katsaka, J.; Habuka, N.; Miyano, M.; Masuta, C.; Koiwai, A.  
 Plant Mol. Biol. 19, 887-889, 1992  
 A:Title: Nucleotide sequence of cDNA encoding beta-luffin, another ribosome-in-  
 A:Reference number: S23519; MUID:92353400; PMID:1643290  
 A:Accession: S23519  
 A:Molecule type: mRNA  
 A:Residues: 1-278 <KAT>  
 A:Cross-references: EMBL:X62372; NID:g19149; PIDN:CAA44230.1; PID:g19150  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 F:26-264/Domain: rRNA N-glycosidase homology <RNG>  
  
 Query Match 30.8%; Score 300.5; DB 2; Length 278;  
 Best Local Similarity 35.3%; Pred. No. 4.7e-20;  
 Matches 66; Conservative 43; Mismatches 57; Indels 21; Gaps 4  
  
 QY 10 INFTTAGATVQSYNNFTRAVRGRL-----TVLPNRVGLPINQRFILVELSNHAEISV 59  
 DB 24 VFSLSGADSKSYSKFITALKRKALPSEKVSNIPLLLPSASGA---SRYLIMQLSNYDAK 80  
 QY 60 SVTLALDVTNAYVVGVRAGNSAYEFHPDNOEDAEATHLFTDVQNRVTFAPGGNYDRLEQ 119  
 DB 81 A:TMALDVTNAYVIMGYLVNSTSYFF---NDSKALASQYVFGKSTIVTLPSYSGNYERLQN 137  
 QY 120 LAGNLRENIELGNGPLBEAISALYYVYTGGTQGLPTLARSPFIICIMISEAARFYIEGEM 179  
 DB 138 AAGKYREKIPLGFRAFDASITSLPHYDS-----TAAAGALFVLIIQTAAESRKYIEGQI 192  
 QY 180 RTRIRYN 186  
 DB 193 IERIPKN 199  
  
 RESULT 13  
 JN0108  
 luffin-b - smooth loofah  
 C:Species: Luffa cylindrica (smooth loofah)  
 C>Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 07-May-1999  
 C:Accession: JN0108  
 R:Islam, M.R.; Hirayama, H.; Funatsu, G.  
 Agric. Biol. Chem. 55, 229-238, 1991  
 A:Title: Complete amino acid sequence of luffin-b, a ribosome-inactivating pro-  
 A:Reference number: JN0108; MUID:91248488; PMID:1368666  
 A:Accession: JN0108  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-250 <ISL>  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 F:5-246/Domain: rRNA N-glycosidase homology <RNG>

QY 10 INFTTAGATVQSYYNFRHVRGRLTVPINRVLGLPIN-----QRFILVELSNHAIE 58  
:  
Db 22 VRFSLSGSSSTYSXKFIQDLR--KALPSN-GTVYNITLLILSSAGSARYTMTLSNYDG 77  
:  
QY 59 LSVTLALDVTWAYVGVYRGASNAFFEHDPNOEDAEATHLFTDVQNRYTFAFGNGNYDRLE 118  
:  
Db 78 KAITVAVDVTVNYIMGYLVNSTISFFF---NESDKLASQYVFKGSTIVLPISGNYEKLQ 134  
:  
QY 119 QLAGNLRENIELGNGPBLEEIASLYXYSTGGTQLTPLARSPFIICTMISEARFYIEGE 178  
:  
Db 135 TAAGKIREKIPLGPPALDSAITLFHYDS----TAAAAAFLVIIQTTAASREFKYIEQG 189  
:  
QY 179 MRTRIRYNR 187  
:  
Db 190 IIERISKNQ 198

RESULT 15  
RUPUG  
RNA N-glycosidase (EC 3.2.2.22) alpha-momorcharin precursor [validated] - balsam pear  
C:Alternate names: agglutinin; momordin-A; ribosome-inactivating protein momorcharin alpi  
C:Species: Momordica charantia (balsam pear, bitter melon)  
C:Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #text\_change 15-Sep-2000  
C:Accession: S14273; A61318; S16490; JN0628; S01670  
R:Hö, W.K.; Liu, S.C.; Shaw, P.C.; Yeung, H.W.; Ng, T.B.; Chan, W.Y.  
Biochim. Biophys. Acta 1088, 311-314, 1991  
A:Title: Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating protein.  
A:Reference number: S14273; MUID:91159486; PMID:2001404  
A:Accession: S14273  
A:Molecule type: mRNA  
A:Residues: 1-286 <ENW>  
A:Cross-references: EMBL:X57682; NID:G19527; PID:CAA040869.1; PID:g19528  
R:Li, S.S.L.  
Experientia 36, 524-527, 1980  
A:Title: Purification and partial characterization of two lectins from Momordica charantia  
A:Reference number: A61318; MUID:80201763; PMID:7379938  
A:Accession: A61318  
A:Molecule type: protein  
A:Residues: 24-50 <LIA>  
A>Note: as a lectin shows agglutinating activity for type-O red blood cells  
R:Montecucchi, P.C.; Lazzerini, A.M.; Barbieri, L.; Stirpe, F.; Sorio, M.; Lippi, D.  
Int. J. Pept. Protein Res. 33, 263-267, 1989  
A:Title: N-terminal sequence of some ribosome-inactivating proteins.  
A:Reference number: S16331; MUID:89326691; PMID:2753596  
A:Accession: S16490  
A:Molecule type: protein  
A:Residues: 24-68, X', 70 <MON>  
R:Minami, Y.; Funatsu, G.  
Biosci. Biotechnol. Biochem. 57, 1141-1144, 1993  
A:Title: The complete amino acid sequence of momordin-a, a ribosome-inactivating protein  
A:Reference number: JN0628; MUID:93372485; PMID:7763984  
A:Accession: JN0628  
A:Molecule type: protein  
A:Residues: 24-107, Q', 109-123, 125-147, 'L', 149-154, 'I', 156-205, 'I', 207-208, 'L', 210-214, 'I'  
R:Experimental source: seed  
submitted to the Brookhaven Protein Data Bank, January 1994  
A:Reference number: A52272; PDB:1AHC  
A:Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 24-269  
R:Husain, J.; Tickie, I.J.; Wood, S.P.  
submitted to the Brookhaven Protein Data Bank, March 1994  
A:Reference number: A52385; PDB:1MOM  
A:Contents: annotation; X-ray crystallography, 2.16 angstroms, residues 24-86, 'L', 88-269  
R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.  
submitted to the Brookhaven Protein Data Bank, July 1994  
A:Reference number: A67089; PDB:1MRH  
A:Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 24-77, 'R', 79-132, 'I'  
C:Function:  
A:Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA there  
C:Superfamily: RNA N-glycosidase; RNA N-glycosidase homology  
C:Keywords: glycoprotein; glycosidase; hydrolase; lectin; seed; toxin  
F1-18/Domain: signal sequence #status predicted <SIG>  
F19-23/Domain: amino-terminal propeptide #status predicted <PRO>

Search completed: February 10, 2004, 16:28:04  
Job time : 11.2817 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:13:55 ; Search time 6.3803 Seconds  
(without alignments)  
1409.756 Million cell updates/sec

Title: US-10-083-336a-11

Perfect score: 975

Sequence: 1 MIFPKQYPIINFITAGATVQ.....RFQVIEGEMTRINRRSA 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	955	97.9	576	1	RICI_RICCO
2	865.5	88.8	564	1	AGGL_RICCO
3	336	34.5	289	1	RIP1_TRIKI
4	327.5	33.6	562	1	ABRC_ABRER
5	326	33.4	289	1	RIPS_TRIKI
6	323	33.1	527	1	ABRB_ABRER
7	320.5	32.9	282	1	RIP2_BRYDI
8	307.5	31.5	563	1	NIGB_SAMNI
9	305	31.3	290	1	RIP1_BRYDI
10	303.5	31.1	528	1	ABRA_ABRER
11	293.5	30.1	250	1	RIPB_LUFCE
12	281.5	28.9	286	1	RIP1_CUCFI
13	274.5	28.2	277	1	RIPA_LUFCE
14	273	28.0	286	1	RIP1_MOMCH
15	272	27.9	286	1	RIP2_MOMBA
16	264	27.1	254	1	MLA_VISAL
17	257	26.4	294	1	RIP1_TRIAN
18	246.5	25.3	316	1	RIPG_GELMU
19	186.5	19.1	294	1	RIPA_PHYAM
20	174	17.8	278	1	RIPP_MIRVA
21	170	17.4	313	1	RIP1_PHYAM
22	168.5	17.3	261	1	RIPS_PHYAM
23	131	13.4	253	1	RIP7_SAPOF
24	127	13.0	253	1	RIP5_SAPOF
25	125	12.8	310	1	RIP2_PHYAM
26	124	12.7	299	1	RIP6_SAPOF
27	121	12.4	292	1	RIP2_SAPOF
28	116.5	11.9	293	1	RIP0_DIAAC
29	114	11.7	280	1	RIP2_HORVU
30	110	11.3	236	1	RIP3_SAPOF
31	108	11.1	319	1	SUTA_BP933
32	107	11.0	280	1	RIP1_HORVU
33	103.5	10.6	300	1	RIP3_MAIZE

34	101.5	10.4	301	1	RIPX_MAIZE
35	97	9.9	304	1	RIP9_MAIZE
36	95	9.7	315	1	SLTA_BPH19
37	94	9.6	315	1	SLTA_BPH30
38	85.5	8.8	1165	1	SYV_AQUAE
39	82	8.4	560	1	J160_HORVU
40	81.5	8.4	305	1	Y348_MYCPN
41	78	8.0	747	1	ORPB_HUMAN
42	76.5	7.8	609	1	HAPT_VIBCH
43	76	7.8	157	1	RIP4_SAPOF
44	75.5	7.7	609	1	NPRV_VIBER
45	75	7.7	224	1	GPH_HAEIN

P28522 zea mays (m)  
P25892 zea mays (m)  
P08026 bacterioph  
P10149 bacterioph  
O67411 aquifex ae  
Q00531 hordeum vul  
P75255 mycoplasma  
Q95xb4 homo sapien  
P24153 vibrio chol  
P27561 saponaria o  
Q00971 vibrio prot  
P44755 haemophilus

## ALIGNMENTS

## RESULT 1

RICI_RICCO	STANDARD;	PRT;	576 AA.
AC	P02879; P02880;		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	13-AUG-1987 (Rel. 05, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Ricin precursor [Contains: Ricin A chain (rRNA N-glycosidase) (EC 3.2.2.22); Ricin B chain].		
OS	Ricinus communis (Castor bean).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Malpighiales; Euphorbiaceae; Ricinus.		
OX	NCBI_TaxID=3988;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86067214; PubMed=2999712;		
RA	Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L., Weaver R.F.;		
RA	"Genomic cloning and characterization of a ricin gene from Ricinus communis.";		
RT	comminis.";		
RL	Nucleic Acids Res. 13:8019-8033(1985).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=92163016; PubMed=1371405;		
RA	Tregear J.W., Roberts L.M.;		
RT	"The lectin gene family of Ricinus communis: cloning of a functional ricin gene and three lectin pseudogenes.";		
RL	Plant Mol. Biol. 18:515-525(1992).		
RN	[3]		
RP	SEQUENCE OF 12-576 FROM N.A.		
RX	MEDLINE=85179479; PubMed=3838723;		
RA	Lamb A., Roberts L.M., Lord J.M.;		
RT	"Nucleotide sequence of cloned cDNA coding for preproricin.";		
RL	Eur. J. Biochem. 148:265-270(1985).		
RN	[4]		
RP	SEQUENCE OF 36-302.		
RX	Yoshitake S., Funatsu G., Funatsu M.;		
RT	"Isolation and sequences of peptic peptides, and the complete sequence of Ile chain of ricin-D.";		
RL	Agric. Biol. Chem. 42:1267-1274(1978).		
RN	[5]		
RP	SEQUENCE OF 315-576.		
RX	Funatsu G., Kimura M., Funatsu M.;		
RT	"Primary structure of Ala chain of ricin D.";		
RL	Agric. Biol. Chem. 43:2221-2224(1979).		
RN	[6]		
RP	CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.		
RX	MEDLINE=90344223; PubMed=1368517;		
RA	Kimura Y., Kusoku H., Tada M., Takagi S., Funatsu G.;		
RT	"Structural analyses of sugar chains from ricin A-chain variant.";		
RL	Agric. Biol. Chem. 54:157-162(1990).		
RN	[7]		
RP	REVIEW.		
RX	MEDLINE=21480122; PubMed=11595634;		
RA	Olsnes S., Kozlov J.V.;		

RT "Ricin."; CC  
 RL Toxicon 39:1723-1728 (2001). CC  
 RN [8] CC  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS). CC  
 RX MEDLINE=87165983; PubMed=3558397; CC  
 RA Monfort W., Villafraanca J.E., Monzingo A.F., Ernst S.R., Katzin B., CC  
 RA Rutenber E., Xiong N.H., Hamlin R., Robertus J.D.; CC  
 RT "The three-dimensional structure of ricin at 2.8 A."; CC  
 RL J. Biol. Chem. 262:5398-5403 (1987). CC  
 RN [9] CC  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN. CC  
 RX MEDLINE=91352004; PubMed=1881881; CC  
 RA Katzin B.J., Collins E.J., Robertus J.D.; CC  
 RT "Structure of ricin A-chain at 2.5 A."; CC  
 RL Proteins 10:251-259 (1991). CC  
 RN [10] CC  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN. CC  
 RX MEDLINE=91352005; PubMed=1881882; CC  
 RA Rutenber E., Robertus J.D.; CC  
 RT "Structure of ricin B-chain at 2.5-A resolution."; CC  
 RL Proteins 10:260-269 (1991). CC  
 RN [11] CC  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN. CC  
 RX MEDLINE=95082010; PubMed=7990130; CC  
 RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J., CC  
 RA Paupit R.A.; CC  
 RT "X-ray structure of recombinant ricin A-chain at 1.8-A resolution."; CC  
 RL J. Mol. Biol. 244:410-422 (1994). CC  
 RN [12] CC  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215. CC  
 RX MEDLINE=96374222; PubMed=8780513; CC  
 RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M., CC  
 RA Molina-Svith M.C., Robertus J.D.; CC  
 RT "Structure and activity of an active site substitution of ricin A CC  
 RT chain."; CC  
 RL Biochemistry 35:11098-11103 (1996). CC  
 RN [13] CC  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN. CC  
 RX MEDLINE=97240820; PubMed=9086280; CC  
 RA Yan X., Hollis T., Svith M., Day P., Monzingo A.F., Malne G.W., CC  
 RA Robertus J.D.; CC  
 RT "Structure-based identification of a ricin inhibitor."; CC  
 RL J. Mol. Biol. 266:1043-1049 (1997). CC  
 RN [14] CC  
 RP MUTAGENESIS. CC  
 RX MEDLINE=93165632; PubMed=1287657; CC  
 RA Kin Y., Robertus J.D.; CC  
 RT "Analysis of several key active site residues of ricin A chain by CC  
 RT mutagenesis and X-ray crystallography."; CC  
 RL Protein Eng. 5:775-779 (1992). CC  
 CC -1- FUNCTION: Ricin is highly toxic to animal cells and to a less CC  
 CC extent to plant cells. The A chain is responsible for inhibiting CC  
 CC protein synthesis through the catalytic inactivation of 60S CC  
 CC ribosomal subunits. It acts as a glycosidase that removes a CC  
 CC specific adenine residue from an exposed loop of 28S ribosomal CC  
 CC RNA. As this loop is involved in the binding of elongation CC  
 CC factors, the modified ribosomes are unable to support protein CC  
 CC synthesis. The A chain can inactivate a few thousand ribosomes CC  
 CC per minute, thus inactivating them faster than the cell can make CC  
 CC new ones. A single A-chain molecule can therefore kill an animal CC  
 CC cell. The B chain binds to cell receptors and facilitates the CC  
 CC entry into the cell of the A chain; B chains are also responsible CC  
 CC for cell agglutination (Lectin activity). It binds to beta-D- CC  
 CC galactopyranoside moieties. CC  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one CC  
 CC specific adenosine on the 28S rRNA. CC  
 CC -1- SUBUNIT: Disulfide-linked dimer of A and B chains. CC  
 CC -1- DOMAIN: The B chain is composed of two domains, each domain CC  
 CC consists of 3 homologous subdomains (alpha, beta, gamma). CC  
 CC -1- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE CC  
 CC MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271. CC  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME- CC  
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY. CC

CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains. CC  
 CC -1- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS CC  
 CC WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3). CC  
 CC -1- DATABASE: NAME=Protein Spotlight; CC  
 CC NOTE=Issue 31 of February 2003; CC  
 CC WWW="http://www.expasy.org/spotlight/articles/sptlt031.html". CC  
 CC ----- CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration CC  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation - CC  
 CC the European Bioinformatics Institute. There are no restrictions on its CC  
 CC use by non-profit institutions as long as its content is in no way CC  
 CC modified and this statement is not removed. Usage by and for commercial CC  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)). CC  
 CC ----- CC  
 CC EMBL; X03179; CAA26939.1; - CC  
 CC EMBL; X52908; CAA37095.1; - CC  
 CC EMBL; X02388; CAA26230.1; - CC  
 CC EMBL; A12892; CAA01058.1; - CC  
 CC FIR; A24041; RLCSD. CC  
 CC PDB; 2AAL; 31-JAN-94. CC  
 CC PDB; 1APG; 31-JAN-94. CC  
 CC PDB; 1FMP; 31-OCT-93. CC  
 CC PDB; 1IFS; 14-JAN-98. CC  
 CC PDB; 1IFT; 14-JAN-98. CC  
 CC PDB; 1IFU; 14-JAN-98. CC  
 CC PDB; 1R7C; 31-OCT-93. CC  
 CC PDB; 1OBS; 16-JUN-97. CC  
 CC PDB; 1OBT; 16-JUN-97. CC  
 CC PDB; 1BR5; 02-SEP-98. CC  
 CC PDB; 1BR6; 02-SEP-98. CC  
 CC PDB; 1IL3; 16-JAN-02. CC  
 CC PDB; 1IL4; 16-JAN-02. CC  
 CC PDB; 1IL9; 16-JAN-02. CC  
 CC GlycoSuiteDB; P02879; - CC  
 CC InterPro; IPR000772; Ricin\_B\_lectin. CC  
 CC InterPro; IPR001574; RIP. CC  
 CC Pfam; PF00652; Ricin\_B\_lectin; 6. CC  
 CC Pfam; PF00161; RIP; 1. CC  
 CC PRINTS; PR00396; SHIGARICIN. CC  
 CC SMART; SM00458; RICIN; 2. CC  
 CC PROSITE; PS00231; RICIN\_B\_LECTIN; 2. CC  
 CC PROSITE; PS00275; SHIGA\_RICIN; 1. CC  
 CC KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; CC  
 CC Glycoprotein; Lectin; Signal; 3D-structure. CC  
 CC FT SIGNAL 1 35 CC  
 CC FT CHAIN 36 302 CC  
 CC FT PEPTIDE 303 314 CC  
 CC FT CHAIN 315 576 CC  
 CC FT DOMAIN 321 448 CC  
 CC FT DOMAIN 451 575 CC  
 CC FT REPEAT 331 373 CC  
 CC FT REPEAT 374 414 CC  
 CC FT REPEAT 417 449 CC  
 CC FT REPEAT 462 497 CC  
 CC FT REPEAT 501 540 CC  
 CC FT REPEAT 543 570 CC  
 CC FT ACT\_SITE 212 212 CC  
 CC FT DISULFID 294 318 CC  
 CC FT DISULFID 334 353 CC  
 CC FT DISULFID 377 394 CC  
 CC FT DISULFID 465 478 CC  
 CC FT DISULFID 504 521 CC  
 CC FT CARBOHYD 45 45 CC  
 CC FT CARBOHYD 271 271 CC  
 CC FT CARBOHYD 409 409 CC  
 CC FT CARBOHYD 449 449 CC  
 CC FT CONFLICT 76 76 CC  
 CC FT CONFLICT 551 551 CC  
 CC FT STRAND 43 47 CC  
 CC FT TURN 49 50 CC

N-LINKED (GLCNAC. . .).  
 /FTIC-CAR 000080.  
 N-LINKED (GLCNAC. . .) (IN MINOR FORM).  
 /FTIC-CAR 000081.  
 N-LINKED (GLCNAC. . .).  
 N-LINKED (GLCNAC. . .).  
 E -> D (IN REF. 3).  
 A -> R (IN REF. 3).



```

Query Match      97.8%; Score 955; DB 1; Length 576;
Best Local Similarity 95.0%; Pred. No. 1.5e-81;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVGRLT-----VLPNRVGLPINQRFILV 51
DB 36 IFPKQYPIINFTTAGATVQSYTNFIRAVGRLTGADVREHIFVLPNRVGLPINQRFILV 95

QY 52 ELSNHAELSVTLALDVNTNAVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111
DB 96 ELSNHAELSVTLALDVNTNAVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 155

QY 112 GNYDRLEQLAGNLRNIELGNGLPBEAISALYYSTGGTQLPPLARSFFIICQMISEAAR 171
DB 156 GNYDRLEQLAGNLRNIELGNGLPBEAISALYYSTGGTQLPPLARSFFIICQMISEAAR 215

QY 172 FOYIEGEMTRIRYNRRA 190
DB 216 FOYIEGEMTRIRYNRRA 234

RESULT 2
AGGL_RICCO STANDARD; PRT; 564 AA.
AC P06750;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DE Agglutinin precursor (RCA) [Contains: Agglutinin A chain (rRNA N-glycosidase) (EC 3.2.2.22); Agglutinin B chain].
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosoids I; Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=86059449; PubMed=2999130;
RA Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.;
RT "The primary sequence of Ricinus communis agglutinin. Comparison with ricin.";
RL J. Biol. Chem. 260:15682-15686(1985).
RN [2]
RN SEQUENCE OF 303-564.
RC TISSUE=Seed;
RA Araki T., Yoshioka Y., Funatsu G.;
RT "The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin isolated from large-grain castor bean seeds.";
RL Biochim. Biophys. Acta 872:277-285(1986).
RN [3]
RN SEQUENCE OF 303-337.
RP MEDLINE=80178723; PubMed=6768555;
RA Lin T.-S., Li S.-L.;
RT "Purification and physicochemical properties of ricins and agglutinins from Ricinus communis.";
RL Eur. J. Biochem. 105:453-459(1980).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M12089; AAA33869.1; -.
CC EMBL; S40368; AAB22584.1; -.

```

```

DR PIR; A24261; RLCSAG.
DR HSPP; P02879; IIR6.
DR GlycosuitedB; P06750; -.
DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP_1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS0231; RICIN_B_LECTIN; 2.
DR PROSITE; PS0275; SHIGA_RICIN; 1.
KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
KW Glycoprotein; Lectin; Signal; POTENTIAL.
FT SIGNAL 1 24
FT CHAIN 25 290 AGGLUTININ A CHAIN.
FT PROPEP 291 302 LINKER PEPTIDE.
FT CHAIN 303 564 AGGLUTININ B CHAIN.
FT DOMAIN 309 436 RICIN B-TYPE LECTIN 1.
FT DOMAIN 439 563 RICIN B-TYPE LECTIN 2.
FT REPEAT 319 361 1-ALPHA.
FT REPEAT 362 402 1-BETA.
FT REPEAT 405 437 1-GAMMA.
FT REPEAT 450 485 2-ALPHA.
FT REPEAT 489 528 2-BETA.
FT REPEAT 531 558 2-GAMMA.
FT ACT_SITE 200 200 BY SIMILARITY.
FT DISULFID 282 306 INTERCHAIN (BY SIMILARITY).
FT DISULFID 322 341 BY SIMILARITY.
FT DISULFID 365 382 BY SIMILARITY.
FT DISULFID 453 466 BY SIMILARITY.
FT DISULFID 492 509 BY SIMILARITY.
FT CARBOHYD 34 34 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 331 331 F -> T (IN REF. 2).
FT CONFLICT 362 362 N -> D (IN REF. 2).
FT CONFLICT 374 374 R -> G (IN REF. 2).
FT CONFLICT 404 404 R -> T (IN REF. 2).
FT CONFLICT 552 552 F -> V (IN REF. 2).
SQ SEQUENCE 564 AA; 62851 MW; D455F2A72F609759 CRC64;

Query Match      88.8%; Score 865.5; DB 1; Length 564;
Best Local Similarity 86.9%; Pred. No. 3.3e-73;
Matches 173; Conservative 7; Mismatches 8; Indels 11; Gaps 2;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVGRLT-----VLPNRVGLPINQRFILV 51
DB 25 IFPKQYPIINFTTAGATVQSYTNFIRAVGRLTGADVREHIFVLPNRVGLPINQRFILV 84

QY 52 ELSNHAELSVTLALDVNTNAVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111
DB 95 ELSNHAELSVTLALDVNTNAVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 144

QY 112 GNYDRLEQLAGNLRNIELGNGLPBEAISALYYSTGGTQLPPLARSFFIICQMISEAAR 171
DB 145 GNYDRLEQLAGNLRNIELGNGLPBEAISALYYSTGGTQLPPLARSFFIICQMISEAAR 203

QY 172 FOYIEGEMTRIRYNRRA 190
DB 204 FOYIEGEMTRIRYNRRA 222

RESULT 3
RIP_TRIKI STANDARD; PRT; 289 AA.
ID RIP_TRIKI
AC P09889;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ribosome-inactivating protein alpha-trichosanthin precursor (rRNA N-glycosidase) (EC 3.2.2.22) (Alpha-TCS).
OS Trichosanthes kirilowii (Mongolian snake-gourd).

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 OX NCBI\_TaxID=3677;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Maximowicz;  
 RX MEDLINE=91153657; PubMed=199291;  
 RA Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;  
 RT "Cloning of trichosanthin cDNA and its expression in *Escherichia*  
 RL *coli*.";  
 RL Gene 97:267-272(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Maximowicz; TISSUE=Leaf;  
 RX MEDLINE=90256790; PubMed=2341400;  
 RA Chow T., Feldman R.A., Lovett M., Piatak M.;  
 RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a  
 RL type I ribosome-inactivating protein.";  
 RL J. Biol. Chem. 265:8670-8674(1990).  
 RN [3]  
 RP SEQUENCE OF 24-270.  
 RC STRAIN=Maximowicz; TISSUE=Tuberos root;  
 RX MEDLINE=90256789; PubMed=2341399;  
 RA Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R.,  
 RA Wu P., Hwang K., Piatak M.;  
 RT "Primary amino acid sequence of alpha-trichosanthin and molecular  
 RL models for abrin A-chain and alpha-trichosanthin.";  
 RL J. Biol. Chem. 265:8665-8669(1990).  
 RN [4]  
 RP SEQUENCE OF 24-270.  
 RC TISSUE=tuberos root;  
 RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,  
 RA Tian G.Y., Ni C.Z.;  
 RT "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and  
 RL application.";  
 RL Pure Appl. Chem. 58:789-798(1986).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).  
 RX MEDLINE=94344957; PubMed=8086085;  
 RA Zhou F., Fu Z., Chen M., Lin Y., Pan K.;  
 RT "Structure of trichosanthin at 1.88-A resolution.";  
 RL Proteins 19:4-13(1994).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).  
 RX MEDLINE=95344383; PubMed=7619070;  
 RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;  
 RT "Studies on crystal structures, active-centre geometry and  
 RL depurinatin mechanism of two ribosome-inactivating proteins.";  
 RL Biochem. J. 309:285-298(1995).  
 CC -!- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS  
 CC CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT  
 CC INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; M34859; AAA34207.1; -;  
 DR EMBL; J05434; AAA34206.1; -;  
 DR PIR; JT0566; RLTT  
 DR PDB; 1MRJ; 07-FEB-95.  
 DR PDB; 1MRK; 07-FEB-95.  
 DR PDB; 1TCS; 10-JUL-95.  
 DR PDB; 1J4G; 28-JAN-03.

DR PDB; 1NLI; 21-JAN-03.  
 DR PDB; IQD2; 24-APR-00.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PRO0396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA RICIN; 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin; Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 270  
 FT PROPEP 271 289  
 FT ACT SITE 183  
 FT CONFLICT 57 60  
 FT CONFLICT 82 84  
 FT CONFLICT 87 87  
 FT CONFLICT 92 92  
 FT CONFLICT 143 144  
 FT CONFLICT 196 196  
 FT CONFLICT 215 216  
 FT CONFLICT 231 231  
 FT CONFLICT 234 234  
 FT CONFLICT 246 266  
 FT CONFLICT 247 247  
 FT STRAND 25 28  
 FT TURN 30 31  
 FT HELIX 34 46  
 FT TURN 47 47  
 FT STRAND 50 54  
 FT TURN 55 56  
 FT STRAND 57 60  
 FT HELIX 66 69  
 FT STRAND 70 76  
 FT TURN 78 79  
 FT STRAND 82 88  
 FT TURN 89 92  
 FT STRAND 93 99  
 FT TURN 100 101  
 FT STRAND 102 105  
 FT HELIX 109 114  
 FT TURN 115 117  
 FT TURN 120 121  
 FT STRAND 124 127  
 FT HELIX 134 141  
 FT TURN 142 142  
 FT HELIX 145 147  
 FT STRAND 150 150  
 FT HELIX 152 163  
 FT TURN 164 165  
 FT HELIX 167 180  
 FT TURN 181 181  
 FT HELIX 182 186  
 FT STRAND 187 187  
 FT HELIX 188 195  
 FT TURN 196 196  
 FT STRAND 202 202  
 FT HELIX 206 226  
 FT TURN 227 230  
 FT STRAND 231 239  
 FT TURN 241 242  
 FT STRAND 245 250  
 FT TURN 251 252  
 FT HELIX 254 258  
 FT TURN 259 259  
 FT STRAND 260 260  
 FT STRAND 263 263  
 FT TURN 266 268  
 FT SEQUENCE 289 AA; 31676 MW; 5CE09BB630575BB9 CRC64;  
 SQ  
 Query Match 34.5%; Score 336; DB 1; Length 289;  
 Best Local Similarity 38.9%; Pred. NO. 3.8e-24;  
 Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;





FT CONFLICT 291 D -> N (IN REF. 2).  
 FT CONFLICT 350 AE -> PQ (IN REF. 2).  
 FT CONFLICT 378 S -> N (IN REF. 2).  
 FT CONFLICT 426 L -> M (IN REF. 2).  
 FT CONFLICT 428 Y -> D (IN REF. 2).  
 FT CONFLICT 431 N -> S (IN REF. 2).  
 FT CONFLICT 431 R -> K (IN REF. 2).  
 FT CONFLICT 484 R -> S (IN REF. 2).  
 FT CONFLICT 491 N -> Y (IN REF. 2).  
 FT CONFLICT 493 H -> G (IN REF. 2).  
 FT CONFLICT 502 E -> Q (IN REF. 2).  
 FT CONFLICT 509 E -> Q (IN REF. 2).  
 FT CONFLICT 513 H -> W (IN REF. 2).  
 FT CONFLICT 516 H -> T (IN REF. 2).  
 SQ SEQUENCE 527 AA; 3253AE49CE9494A CRC64;  
 Query Match 33.1%; Score 323; DB 1; Length 527;  
 Best Local Similarity 43.9%; Pred. No. 1.3e-22;  
 Matches 83; Conservative 21; Mismatches 69; Indels 16; Gaps 4;  
 QY 6 QYPLINFTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFILVELSNHA 57  
 Db 1 QDQVKEITTEGATSSQYKQFIEALRQLTGGLIHGIPVLPDPTTLQERNYISVELSNSD 60  
 QY 58 ELSTALDVTNAYVGYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAPGNYDRL 117  
 Db 61 TESTEAGIDVSNAYVAYRAGNSRYFL---RDAPTSASRYLFTGTQ-QYSLRFNGSYIDL 116  
 QY 118 EQLAGNLENIELGNGLPLEEASLYYSTGGTQLPTLARSFFICIONISEARFOYIE 177  
 Db 117 ERLARQTRQIPLGLQALRHAIISFL-----QSGTDDQEIARTLVIQWASEAARYRFSY 172  
 QY 178 EMRTIRYN 186  
 Db 173 RVGSIRTN 181  
 RESULT 7  
 ID RIP2 BRYDI STANDARD; PRT; 282 AA.  
 AC P98184; Q98J0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein bryodin II precursor (rRNA N-glycosidase) (EC 3.2.2.22) (BD2).  
 OS Bryonia dioica (Red bryony).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.  
 OX NCBI\_TaxID=3652;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Slegall C.B., Gawlak S.L., Marquardt H.;  
 RT "Bryodin 2 a ribosome-inactivating protein isolated from the plant Bryonia dioica.";  
 RL Patent number US5597569, 28-JAN-1997.  
 RN [2]  
 RP SEQUENCE OF 22-42.  
 RC TISSUE=Root;  
 RX MEDLINE=95151812; PubMed=7849072;  
 RA Slegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,  
 RA Marquardt H.;  
 RT "Characterization of ribosome-inactivating proteins isolated from Bryonia dioica and their utility as carcinoma-reactive immunoconjugates.";  
 RL Bioconj. Chem. 5:423-429(1994).  
 CC -!- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 1 RIP SUBFAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; I34238; ; NOT\_ANNOTATED\_CDS.  
 DR HSP; P09989; IMRJ.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KM Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;  
 KM Multigene family; Glycoprotein; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 282 RIBOSOME-INACTIVATING PROTEIN BRYODIN II.  
 FT ACT SITE 183 183 BY SIMILARITY.  
 FT CARBOHYD 25 25 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 282 AA; 30754 MW; C52BE2F6A873769C CRC64;  
 Query Match 32.9%; Score 320.5; DB 1; Length 282;  
 Best Local Similarity 46.0%; Pred. No. 1e-22;  
 Matches 81; Conservative 24; Mismatches 54; Indels 17; Gaps 7;  
 QY 10 INFTTAGATVQSYTNFIRAVRGRLT-LPNRVGLPINQ-----RFLVELSNHAELSV 61  
 Db 24 INFSLGATGATYKTRIRNLTKLTGTGTPRVYDIPVRNAAAGLARFQLVTLTNGESV 83  
 QY 62 TLALDVTNAYVGYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAPGNYDRLQLA 121  
 Db 84 TVALDVTNAYVGYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAPGNYDRLQLA 139  
 QY 122 GNL-RENIELGNGLPLEEASLYYSTGGTQLPTLARSFFICIONISEARFOYIE 176  
 Db 140 GRISRENIELGFSEISSAIGNMFRNP-GTSVP---RAFTVIQTVSEARFKYIE 191  
 RESULT 8  
 ID NIGB SAMNI STANDARD; PRT; 563 AA.  
 AC P33183; P33184; P93542;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Nigrin b precursor (Agglutinin V) (SNAV) [Contains: Nigrin b A chain (rRNA N-glycosidase) (EC 3.2.2.22); Nigrin b B chain].  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; Campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Barb;  
 RX MEDLINE=96215449; PubMed=8647092;  
 RA Van Damme E.J., Barre A., Rouge P., Van Leuven E., Peumans W.J.;  
 RT "Characterization and molecular cloning of Sambucus nigra agglutinin V (nigrin b), a GalNAC-specific type-2 ribosome-inactivating protein from the bark of elderberry (Sambucus nigra).";  
 RL Eur. J. Biochem. 237:505-513(1996).  
 RN [2]  
 RP SEQUENCE OF 26-49 AND 298-321.  
 RC TISSUE=Barb;  
 RX MEDLINE=94003077; PubMed=8400135;  
 RA Girbes T., Citores L., Ferreras J.M., Rojo M.A., Iglesias R., Munoz R., Arias F.J., Calonge M., Garcia J.R., Mendez E.;  
 RT "Isolation and partial characterization of nigrin b, a non-toxic novel type 2 ribosome-inactivating protein from the bark of Sambucus nigra L.";  
 RL Eur. J. Biochem. 237:505-513(1996).  
 CC -!- FUNCTION: NON-TOXIC TYPE 2 RIP WHICH STRONGLY INHIBITS MAMMALIAN





CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M98344; AAA32624.1; ALT INIT.  
CC EMBL; X54872; -; NOT ANNOTATED\_CDS.  
CC PIR; S32429; TZLSA.  
CC PDB; 1ABR; 07-FEB-95.  
CC InterPro; IPR000772; Ricin\_B\_lectin.  
CC InterPro; IPR001574; RIP.  
CC Pfam; PF00852; Ricin\_B\_lectin; 6.  
CC Pfam; PF00161; RIP; 1.  
CC PRINTS; PR00396; SHIGARICIN.  
CC SMART; SM00458; RICIN; 2.  
CC PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
CC PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
KW Glycoprotein; Lectin; 3D-structure; Pyrrolidone carboxylic acid.  
FT CHAIN 1 251 ABRIN-A A CHAIN.  
FT PEPTIDE 252 261 LINKER PEPTIDE.  
FT CHAIN 262 528 ABRIN-A B CHAIN.  
FT DOMAIN 273 400 RICIN B-TYPE LECTIN 1.  
FT DOMAIN 403 527 RICIN B-TYPE LECTIN 2.  
FT REPEAT 283 325 1-ALPHA.  
FT REPEAT 326 366 1-BETA.  
FT REPEAT 369 401 1-GAMMA.  
FT REPEAT 414 449 2-ALPHA.  
FT REPEAT 453 492 2-BETA.  
FT REPEAT 495 528 2-GAMMA.  
FT ACT\_SITE 164 164 BY SIMILARITY.  
FT DISULFID 247 269 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 286 305 BY SIMILARITY.  
FT DISULFID 329 346 BY SIMILARITY.  
FT DISULFID 417 430 BY SIMILARITY.  
FT DISULFID 456 473 BY SIMILARITY.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .).  
FT CONFLICT 202 202 MISSING (IN REF. 2).  
FT CONFLICT 298 298 N -> Y (IN REF. 4).  
FT CONFLICT 427 427 M -> L (IN REF. 4).  
FT CONFLICT 467 467 T -> P (IN REF. 4).  
FT CONFLICT 483 483 V -> L (IN REF. 4).  
FT STRAND 5 8  
FT TURN 10 11  
FT HELIX 14 28  
FT STRAND 32 33  
FT TURN 34 35  
FT STRAND 36 38  
FT TURN 42 43  
FT HELIX 47 49  
FT STRAND 51 57  
FT STRAND 63 69  
FT TURN 70 72  
FT STRAND 75 79  
FT STRAND 83 86  
FT TURN 88 89  
FT TURN 92 93  
FT HELIX 94 97  
FT TURN 100 101  
FT STRAND 103 106  
FT TURN 113 114  
FT HELIX 115 119  
FT TURN 124 126  
FT STRAND 129 131  
FT HELIX 131 142

FT TURN 143 144  
FT HELIX 148 167  
FT STRAND 168 168  
FT TURN 169 180  
FT STRAND 181 182  
FT STRAND 185 185  
FT HELIX 189 196  
FT TURN 197 197  
FT HELIX 198 207  
FT STRAND 212 220  
FT TURN 222 223  
FT STRAND 226 231  
FT TURN 232 233  
FT HELIX 235 239  
FT STRAND 240 240  
FT STRAND 243 243  
FT STRAND 248 248  
FT STRAND 268 268  
FT STRAND 276 277  
FT STRAND 279 280  
FT HELIX 282 284  
FT STRAND 286 289  
FT HELIX 290 292  
FT TURN 296 297  
FT STRAND 299 303  
FT HELIX 311 313  
FT STRAND 315 317  
FT TURN 319 320  
FT STRAND 322 325  
FT TURN 326 327  
FT STRAND 328 332  
FT TURN 337 338  
FT STRAND 340 344  
FT TURN 346 348  
FT HELIX 351 353  
FT STRAND 355 355  
FT STRAND 357 358  
FT TURN 360 361  
FT STRAND 364 366  
FT TURN 367 370  
FT STRAND 371 374  
FT TURN 380 381  
FT STRAND 383 383  
FT STRAND 385 387  
FT HELIX 393 395  
FT STRAND 398 399  
FT STRAND 406 408  
FT STRAND 410 411  
FT HELIX 413 415  
FT STRAND 417 421  
FT TURN 422 423  
FT STRAND 424 428  
FT TURN 432 433  
FT HELIX 435 437  
FT STRAND 439 441  
FT TURN 443 444  
FT STRAND 447 449  
FT TURN 450 451  
FT STRAND 452 459  
FT TURN 464 465  
FT STRAND 467 472  
FT TURN 474 475  
FT HELIX 478 480  
FT STRAND 483 484  
FT TURN 486 487  
FT STRAND 490 492  
FT TURN 493 496  
FT STRAND 497 501

Query Match 31.1%; Score 303.5; DB 1; Length 528;  
Best Local Similarity 41.1%; Pred. No. 8.6e-21;  
Matches 76; Conservative 26; Mismatches 64; Indels 19; Gaps 5;



```

Query Match      30.1%; Score 293.5; DB 1; Length 250;
Best Local Similarity 33.7%; Pred. No. 3e-20;
Matches 63; Conservative 46; Mismatches 57; Indels 21; Gaps 4

QY      10  INFTAGATGVSQYTNFIRAVRGRL-----TVLPNRVGGPINRQIRLVELSNHAEL 59
Db      3  VSFSLSGADSKSYSKFTALRKALPSEKVSNTPLLPSASGA-----SRVLMQLSNYDAK 59
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      60  SVTLALDVNYVGYRAGNSAYFHPDNQEDAEATLHFTVQNRYTFAFGNYDRLEQ 119
Db      60  AITMALDVNYVIMGYLVNSTSYF---ANESDKLASQVYFKGSTLVTVPYSGNYERLQN 116
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      120  LAGNLRNEILGNGLPEEAISALYSYTGGTQLPTLARSFICIOMISAARFQYIEGEM 179
Db      117  ARGKIREKIPLGFRALDSALTSIFHYDS-----TAAAAAFLVIQTTAASRFKYIEGI 171
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      180  RTRIRYN 186
Db      172  IERIPKN 178

```

Query Match	28.9%;	Score	281.5;	DB 1;	Length	286;			
Best Local Similarity	32.4%;	Pred. No.	4.6e-19;						
Matches	60;	Conservative	46;	Mismatches	64;	Indels	15;	Gaps	3;

  

QY	10	INFTTAGATGQSYNETRAVGRGLTIVLPNRVGLPI-----NQRFILVELSNHAELSVT	62
DB	28	VKFSLLGNHKSXKFTSMRNALPNAGDITYNPLLVPSISGRRYILMQLSNYGNITIT	87
QY	63	LALDVTNAYVVGVRAGNSAYFFHPFDNQDEAAETHLFTDVQNRYYTFAFGNNTDRLEQLAG	122
DB	88	MAVDVTNVIYIMGLVNGTSYFF--NETDAQLASKEVFQGTKSITLTPYSGNYOKLQSVAR	144
QY	123	NLRNIEGLNGPLEEATSAIYYVSTGGTQLPTLARSFIICIQISENARFPVQIEGEMETR	182
DB	145	KERDSIPGFWMALSAISTLIYYDDRSAPI-----AFLVLIOTTAPAAKYKVIKQIIDR	199
QY	183	IRYNR	187
DB	200	ISVSK	204

```

RIPA_LUFFCY          STANDARD;          PRT;      277 AA.
ID   Q00456;
AC   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DE   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ribosome-inactivating protein luffin-alpha precursor (rRNA
DE   N-glycosidase) (EC 3.2.2.22).
OS   Luffa cylindrica (Smooth loofah) (Sponge gourd).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Luffa.
OX   NCBI_TaxID=3670;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Seed;
RX   MEDLINE=92288316; PubMed=1600156;
RA   Kataoka J., Habuka N., Miyano M., Masuta C., Koiwai A.;
RT   "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-
RT   inactivating protein from Luffa cylindrica.";
RL   Plant Mol. Biol. 18:1199-1202(1992).
CC   -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC   specific adenosine on the 28S rRNA.
CC   -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC   TYPE 1 RIP SUBFAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC   or send an email to license@isb-sib.ch).
CC   -----
CC   EMBL: X62371; CAA44229.1; --
CC   DR   PIR: S22494; S22494.
CC   DR   HSP; P16094; IAHG.
CC   DR   InterPro: IPR001574; RIP.
CC   DR   Pfam: PF00161; RIP; 1.
CC   DR   PRINTS; PR00396; SHIGARICIN.
CC   DR   PROSITE; PS00275; SHIGA_RICIN; 1.
CC   KW   Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
CC   KW   Toxin; Signal.
CC   FT   SIGNAL 1 19
CC   FT   CHAIN 20 277
CC   FT   ACT SITE 179 179
CC   FT   BY SIMILARITY.
CC   FT   RIBOSOME-INACTIVATING PROTEIN LUFFIN-
CC   FT   ALPHA.
CC   SQ   SEQUENCE 277 AA; 30212 MW; EA17FC27998C25AC CRC64;
CC   Query Match 28.2%; Score 274.5; DB 1; Length 277;
CC   Best Local Similarity 33.3%; Pred. No. 2e-18;
CC   Matches 63; Conservative 42; Mismatches 61; Indels 23; Gaps 5;
CC   -----
QY   10 INFTAGATVQSYNFTFRVAVRGLTLPNVRVGLPIN-----QRFILVELSNHAE 58
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   22 VRSLSSSSSTSKFKGCDLR---KALPSN-GTYNTLLSSASGRITLMLSNYDG 77
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   59 LSVFLADVTNAYVGVYAGNRSAYFFHPDQEDAEATHLFTDQVNRYYTFAFGNGYDRLE 118
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   78 KAITVAVDVTNVIYMGVLVNSTSYFF---NESDAKLASQYVFKGSTIVTLPYSGNYEKLQ 134
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   119 QLGNLRNLELNGPLEEALISALYYSTGTQTLPTLARSFICQIMSEARFQYIEGE 178
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   135 TAAGKIREKIPLGSPALDSAITLTFHYDS-----TAAAAAFLVIQTAAASRKYIEGQ 189
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   179 MRTRIRYNR 187
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   190 IIERISKQ 198
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 14

RIP1\_MOMCH

ID\_RIP1\_MOMCH STANDARD; PRT; 286 AA.

```

AC   P16094; P24697;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ribosome-inactivating protein momordin I precursor (rRNA
DE   N-glycosidase) (EC 3.2.2.22) (Alpha-momorcharin) (Alpha-MMC).
OS   Momordica charantia (Bitter melon) (Balsam pear).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX   NCBI_TaxID=3673;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Seed;
RX   MEDLINE=91159486; PubMed=2001404;
RA   Ho W.K.K., Liu S.C., Shaw P.C., Yeung H.W., Ng T.B., Chan W.Y.;
RT   "Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating
RT   protein.";
RL   Biochim. Biophys. Acta 1088:311-314(1991).
CC   [2]
CC   SEQUENCE OF 24-38.
CC   TISSUE=Seed;
CC   RX   MEDLINE=89326691; PubMed=2753596;
CC   RA   Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,
CC   RA   Lappi D.;
CC   RT   "N-terminal sequence of some ribosome-inactivating proteins.";
CC   RT   Int. J. Pept. Protein Res. 33:263-267(1989).
CC   [3]
CC   SEQUENCE OF 24-70.
CC   TISSUE=Seed;
CC   RX   MEDLINE=89005108; PubMed=3262509;
CC   RA   Casellas P., Dussosoy D., Falasca A.I., Barbieri L., Stirpe F.;
CC   RA   Guillemot J.C., Ferrara P., Bolognesi A., Cennini P.,
CC   RA   Trichokirin, a ribosome-inactivating protein from the seeds of
CC   RT   Trichosanthes kirilowii Maximowicz. Purification, partial
CC   RT   characterization and use for preparation of immunotoxins.";
CC   RT   Eur. J. Biochem. 176:581-588(1988).
CC   [4]
CC   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
CC   RP   MEDLINE=94356447; PubMed=8075985;
CC   RA   Ren J., Wang Y., Dong Y., Stuart D.I.;
CC   RT   "The N-glycosidase mechanism of ribosome-inactivating proteins
CC   RT   implied by crystal structures of alpha-momorcharin.";
CC   RT   Structure 2:7-16(1994).
CC   [5]
CC   X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).
CC   RP   MEDLINE=94192822; PubMed=8143869;
CC   RA   Husain J., Tickle I.J., Wood S.P.;
CC   RT   "Crystal structure of momordin, a type I ribosome inactivating
CC   RT   protein from the seeds of Momordica charantia.";
CC   RT   FEBS Lett. 342:154-158(1994).
CC   [6]
CC   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
CC   RP   MEDLINE=95344383; PubMed=7619070;
CC   RA   Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
CC   RT   "Studies on crystal structures, active-centre geometry and
CC   RT   depurinating mechanism of two ribosome-inactivating proteins.";
CC   RT   Biochem. J. 309:285-298(1995).
CC   -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC   specific adenosine on the 28S rRNA.
CC   -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC   TYPE 1 RIP SUBFAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC   or send an email to license@isb-sib.ch).
CC   -----
CC   EMBL: X57682; CAA40869.1; --
CC   DR   PIR: S14273; RUPUGG.

```



Db 141 KIRENIDLGPLSSAITTLFYNA-----QSAPSAIIVLIQTTAAARFKYIERHVAKY 195

QY 183 IRYN 186

Db : |

196 VATN 199

Search completed: February 10, 2004, 16:23:26  
Job time : 6.33803 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:00 ; Search time 25.3521 Seconds  
(without alignments)  
1933.961 Million cell updates/sec

Title: US-10-083-336A-11

Perfect score: 975

Sequence: 1 MFPPQYPIINTTAGATVQ.....RFQYIEGEMRTIRYNRRA 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:\*

- 1: sp archaea.\*
- 2: sp bacteria.\*
- 3: sp fungi.\*
- 4: sp human.\*
- 5: sp invertebrate.\*
- 6: sp mammal.\*
- 7: sp mhc.\*
- 8: sp organelle.\*
- 9: sp phage.\*
- 10: sp plant.\*
- 11: sp rodent.\*
- 12: sp virus.\*
- 13: sp vertebrate.\*
- 14: sp unclassified.\*
- 15: sp virus.\*
- 16: sp bacteriap.\*
- 17: sp archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	955	97.9	541	10	Q41174
2	374.5	38.4	580	10	Q94BW3
3	372.5	38.2	580	10	Q94BW4
4	370.5	38.0	581	10	Q94BW5
5	368.5	37.8	549	10	Q94BV22
6	334	34.3	289	10	Q94KE4
7	332	34.1	289	10	Q41216
8	329.5	33.8	563	10	Q04367
9	326	33.4	247	10	Q19LRE3
10	323.5	33.2	564	10	Q94VR2
11	318.5	32.7	528	10	Q6076
12	317	32.5	270	10	Q81PV7
13	310.5	31.8	252	10	Q38760
14	307.5	31.5	563	10	Q945S2
15	307.5	31.5	563	10	Q8GT32
16	305	31.3	270	10	Q41611

17	303.5	31.1	252	10	Q38761
18	302	31.0	565	10	Q04071
19	300.5	30.8	278	10	Q00980
20	298.5	30.6	547	10	Q9M6E9
21	297.5	30.5	251	10	Q96236
22	296.5	30.4	251	10	Q96237
23	293.5	30.1	566	10	Q04072
24	290.5	29.8	251	10	Q96235
25	273	28.0	264	10	Q9FSH2
26	273	28.0	570	10	Q41358
27	271	27.8	592	10	Q8W2E7
28	270	27.7	249	10	Q8LKQ5
29	270	27.7	286	10	Q9FUW7
30	267	27.4	570	10	Q22415
31	266.5	27.3	604	10	Q9M654
32	265	27.2	286	10	Q41257
33	263	27.0	565	10	Q8W243
34	260	26.7	254	10	Q8LKQ6
35	260	26.7	573	10	Q8W2E8
36	259	26.6	251	10	Q8LKQ4
37	259	26.6	293	10	Q8S452
38	259	26.6	569	10	Q93543
39	253	25.9	531	10	Q8RXH6
40	251	25.7	249	10	Q8RXH7
41	250	25.6	258	10	Q9S9E4
42	247	25.3	293	10	Q8VYU0
43	246.5	25.3	203	10	Q8RY69
44	246.5	25.3	275	10	Q8H1Y4
45	245.5	25.2	603	10	Q9M653

#### ALIGNMENTS

#### RESULT 1

Q41174 ID Q41174 PRELIMINARY; PRT; 541 AA.  
AC Q41174;  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)  
DE Proridin A chain (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).  
DE Ricinus communis (Castor bean).  
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
OX NCBI\_TaxID=3988;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92338377; PubMed=1633311;  
RA Roberts L.M., Tregear J.W., Lord J.M.;  
RT "Molecular cloning of ricin."  
RL Targeted Diagn. Ther. 7:81-97(1992).  
CC -|- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -|- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; S40366; AAB22582.1; -.  
DR HSSP; P02879; 1BR6.  
DR InterPro; IPR000772; Ricin\_B\_lectin.  
DR InterPro; IPR001574; RIP.  
DR InterPro; IPR001400; Somatotropin.  
DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; P000396; SHIGARICIN.  
DR SMART; SM00458; RICIN; 2.  
DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
KW Hydrolase; Toxin.  
FT NON TER 1  
SQ SEQUENCE 541 AA; 60281 MW; 2B7B2CDEF1F2E9D9 CRC64;

Q38761	abrus preca
Q04071	sambucus ni
Q00980	luffa cylin
Q9M6E9	abrus preca
Q96236	abrus preca
Q96237	abrus preca
Q04072	sambucus ni
Q96235	abrus preca
Q9FSH2	momordica c
Q41358	sambucus ni
Q8W2E7	iris hollan
Q8LKQ5	viscum albu
Q9FUW7	momordica c
Q22415	sambucus ni
Q9M654	polygonatum
Q41257	momordica c
Q8W243	viscum albu
Q8LKQ6	viscum albu
Q8W2E8	iris hollan
Q8LKQ4	viscum albu
Q8S452	jatropha cu
P93543	sambucus ni
Q8RXH6	viscum albu
Q8RXH7	viscum albu
Q9S9E4	gelonium mu
Q8VYU0	jatropha cu
Q8RY69	gynostemma
Q8H1Y4	polygonatum
Q9M653	polygonatum

```
Query Match          97.9%; Score 955; DB 10; Length 541;
Best Local Similarity 95.0%; Pred.No.1.3e-82;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVGRLT-----VLPNRVGLPINQRFILV 51
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVGRLTGADVRHDIPLVLPNRVGLPINQRFILV 60

QY 52 ELSNHAELSVTLALDVTVNAVYVGYRAGNSAYFFHPDNOQDAEAITHLFTDVQNRVTFAPG 111
DB 61 ELSNHAELSVTLALDVTVNAVYVGYRAGNSAYFFHPDNOQDAEAITHLFTDVQNRVTFAPG 120

QY 112 GNYDRLEQLAGNLRENELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAAR 171
DB 121 GNYDRLEQLAGNLRENELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAAR 180

QY 172 FQYIEGEMTRINRRSA 190
DB 181 FQYIEGEMTRINRRSA 199

RESULT 2
Q94BW3 PRELIMINARY; PRT; 580 AA.
ID Q94BW3;
AC Q94BW3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type 2 ribosome-inactivating protein cinnamomin III precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q.; Gong Z.Z.; Liu W.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomin proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC EMBL; AY039803; AAK82460.1; -
CC InterPro; IPR000772; Ricin_B_lectin.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00652; Ricin_B_lectin; 6.
CC Pfam; PF00161; RIP; 1.
CC PRINTS; PR00396; SHIGARICIN.
CC SMART; SM00458; RICIN; 2.
CC PROSITE; PS0231; RICIN_B_LECTIN; 2.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 32
FT CHAIN 33 580
FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOMIN III.
SQ SEQUENCE 580 AA; 64421 MW; 940D10F01E7FB558 CRC64;

Query Match          38.4%; Score 374.5; DB 10; Length 580;
Best Local Similarity 48.1%; Pred.No.3.9e-27;
Matches 90; Conservative 26; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFTTAGATVQSYTNFIRAVGRLT-----VLPNRVGLPINQRFILVLSN-H 56
DB 33 YQVTVTFTTKNATKTSYTFEIALRAQLASGEHPHGPVVRERSTVPDSKRFILVLSNWA 92

QY 57 AELSVTLALDVTVNAVYVGYRAGNSAYFFHPDNOQDAEAITHLFTDVQNRVTFAPGNYDR 116
DB 93 ADSPVTLAVDVTVNAVYVGYRAGNSAYFFHPDNOQDAEAITHLFTDVQNRVTFAPGNYDR 149

QY 117 LEQAGNLRENELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAARFOYTE 176
DB 150 LERVAGEREELLGMDPLENAISALWISNL--NQORALARSLIVVIQWVAEAVRFRFIE 207

Query Match          38.2%; Score 372.5; DB 10; Length 580;
Best Local Similarity 47.6%; Pred.No.6.1e-27;
Matches 89; Conservative 27; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFTTAGATVQSYTNFIRAVGRLT-----VLPNRVGLPINQRFILVLSN-H 56
DB 33 YQVTVTFTTKNATKTSYTFEIALRAQLASGEHPHGPVVRERSTVPDSKRFILVLSNWA 92

QY 57 AELSVTLALDVTVNAVYVGYRAGNSAYFFHPDNOQDAEAITHLFTDVQNRVTFAPGNYDR 116
DB 93 ADSPVTLAVDVTVNAVYVGYRAGNSAYFFHPDNOQDAEAITHLFTDVQNRVTFAPGNYDR 149

QY 117 LEQAGNLRENELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAARFOYTE 176
DB 150 LERVAGEREELLGMDPLENAISALWISNL--NQORALARSLIVVIQWVAEAVRFRFIE 207

Query Match          38.4%; Score 374.5; DB 10; Length 580;
Best Local Similarity 48.1%; Pred.No.3.9e-27;
Matches 90; Conservative 26; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFTTAGATVQSYTNFIRAVGRLT-----VLPNRVGLPINQRFILVLSN-H 56
DB 33 YQVTVTFTTKNATKTSYTFEIALRAQLASGEHPHGPVVRERSTVPDSKRFILVLSNWA 92

QY 57 AELSVTLALDVTVNAVYVGYRAGNSAYFFHPDNOQDAEAITHLFTDVQNRVTFAPGNYDR 116
DB 93 ADSPVTLAVDVTVNAVYVGYRAGNSAYFFHPDNOQDAEAITHLFTDVQNRVTFAPGNYDR 149

QY 117 LEQAGNLRENELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAARFOYTE 176
DB 150 LERVAGEREELLGMDPLENAISALWISNL--NQORALARSLIVVIQWVAEAVRFRFIE 207
```

```
QY 177 GEMTRIR 183
DB 208 YRVRESI 214

RESULT 3
Q94BW4 PRELIMINARY; PRT; 580 AA.
ID Q94BW4;
AC Q94BW4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type 2 ribosome-inactivating protein cinnamomin II precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q.; Gong Z.Z.; Liu W.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomin proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC EMBL; AY039802; AAK82459.1; -
CC InterPro; IPR000772; Ricin_B_lectin.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00652; Ricin_B_lectin; 6.
CC Pfam; PF00161; RIP; 1.
CC PRINTS; PR00396; SHIGARICIN.
CC SMART; SM00458; RICIN; 2.
CC PROSITE; PS0231; RICIN_B_LECTIN; 2.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 32
FT CHAIN 33 580
FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOMIN II.
SQ SEQUENCE 580 AA; 64265 MW; 37E4289ECCE0CBFF CRC64;

Query Match          38.2%; Score 372.5; DB 10; Length 580;
Best Local Similarity 47.6%; Pred.No.6.1e-27;
Matches 89; Conservative 27; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFTTAGATVQSYTNFIRAVGRLT-----VLPNRVGLPINQRFILVLSN-H 56
DB 33 YQVTVTFTTKNATKTSYTFEIALRAQLASGEHPHGPVVRERSTVPDSKRFILVLSNWA 92

QY 57 AELSVTLALDVTVNAVYVGYRAGNSAYFFHPDNOQDAEAITHLFTDVQNRVTFAPGNYDR 116
DB 93 ADSPVTLAVDVTVNAVYVGYRAGNSAYFFHPDNOQDAEAITHLFTDVQNRVTFAPGNYDR 149

QY 117 LEQAGNLRENELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAARFOYTE 176
DB 150 LERVAGEREELLGMDPLENAISALWISNL--NQORALARSLIVVIQWVAEAVRFRFIE 207

Query Match          38.4%; Score 374.5; DB 10; Length 580;
Best Local Similarity 48.1%; Pred.No.3.9e-27;
Matches 90; Conservative 26; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFTTAGATVQSYTNFIRAVGRLT-----VLPNRVGLPINQRFILVLSN-H 56
DB 33 YQVTVTFTTKNATKTSYTFEIALRAQLASGEHPHGPVVRERSTVPDSKRFILVLSNWA 92

QY 57 AELSVTLALDVTVNAVYVGYRAGNSAYFFHPDNOQDAEAITHLFTDVQNRVTFAPGNYDR 116
DB 93 ADSPVTLAVDVTVNAVYVGYRAGNSAYFFHPDNOQDAEAITHLFTDVQNRVTFAPGNYDR 149

QY 117 LEQAGNLRENELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAARFOYTE 176
DB 150 LERVAGEREELLGMDPLENAISALWISNL--NQORALARSLIVVIQWVAEAVRFRFIE 207
```

```
RESULT 4
Q94BW5 PRELIMINARY; PRT; 581 AA.
ID Q94BW5;
AC Q94BW5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type 2 ribosome-inactivating protein cinnamomin I precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase).
```

RESULT 5	
Q9FV22	
ID	PRELIMINARY; PRT; 549 AA.
AC	Q9FV22;
DT	01-WAR-2001 (TrEMBLrel. 16, Created)
DT	01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE	Type II ribosome-inactivating protein cinnamomin (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).
DE	Cinnamomum camphora (Camphor tree).
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX	[NCBI_TaxID=13429;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Xie L., Liu W.-Y., Wang E.-D.;
RT	"Molecular cloning of cinnamomin A-, B-chain and the expression, purification, characterization and mutagenesis of the A-Chain.";
RL	Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
CC	-!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE SPECIFIC ADENOSINE ON THE 28S RRNA.
CC	-!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR	EMBL; AF259548; AAF68978.2; -.
DR	HSSP; P02879; 2AAI.

	Query Match	34.3%	Score	334;	DB	10;	Length	289;	
	Best Local Similarity	38.9%;	Pred. No.	1.1e-23;					
	Matches	72;	Conservative	42;	Mismatches	49;	Indels	22; Gaps	5;
QY	10	INFTTAGATVQSNTFIRAVRGRLTVLPN-----RVGVPINORFTLVLSNHAEI	59	:::	:	:	:	:	:
Dd	25	VSPFLSGATSSSVCVFISNR---KALPNRKUVDPLRSSUGGYRALIHITVADE	81	:::	:	:	:	:	:
QY	60	SVTVALDVTNAVVGVRAGNSAFYFFHPDNQEDA-EAITHLFTDVQMYRTFAFGCNYDRLE	118	:::::	:	:	:	:	:





```

SQ      SEQUENCE          564 AA;  62694 MW;  8261681A6DB55CB8 CRC64;
Query Match              33.2%;   Score 323.5;   DB 10;   Length 564;
Best Local Similarity    40.0%;   Pred. No. 2.8e-22;
Matches       76; Conservative 38; Mismatches 51; Indels 25; Gaps 6;

QY      7 YPIINFTTAGATVQSNTNFIRAVGR-----LTVLNPNRVGLPINORFILVELSNH 56
Db            :|::|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
QY      28 YPSVSFNLAGAKSTTYRDFLKNLRDRAVGATGYEVNGLPVLRRESEVOVKRFVLRITNY 87
Db            :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
QY      57 AELSVTLALDVTVNAVYVGVTRAGNSAYFFHPDNQEDAEAI--THLFTDVQNRYTFAGSNI 114
Db            :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
QY      88 NGDTVTSAVDVNTNLXLYAFSANSGSYF----KDATELOKSMLFLT-TQTHTLSFTGNY 141
Db            :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
QY     115 DRLEQLAGLNRENIELGNGLPDEIAISALYYYSGGTGLPTLASFLICIMISAARFOY 174
Db            :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
QY     142 DNLETAAGTRRESIEIGPPNPDLCAITSWY--DGG-----VARSLIVLIQKWPEAAFRY 194
Db            :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
QY     175 IEGEMRTIR 184
Db           ||:|:|:~
QY     195 IEQEVRRLQ 204
Db           ||:|:|:~

RESULT 11
QO6076 PRELIMINARY; PRT; 528 AA.
ID QO6076;
AC QO6076;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Abrin-d (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosoids I; Fabales; Fabaceae; Papilionoideae; Abrease; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93132798; PubMed=8421313;
RT Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
RT "Primary structure of three distinct isoabirins determined by cDNA
RT sequencing: conservation and significance.";
RT J. Mol. Biol. 229:263-267(1993).
CC -!- CATALYTIC ACTIVITY: ENDOPHYLOSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CCC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; M98346; AAA32626.1; -.
DR HSPB; F11140; LABR.
DR InterPro; IPRO00772; Ricin_B_lectin.
DR InterPro; IPRO01574; RIP.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS00231; RICIN_B_LECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
DK Hydrolase; Toxin.
KW NON TER 1
FT NON TER 528
FT SEQUENCE 528 AA; 58870 MW; 62ED42FB8FFE60F8 CRC64;

Query Match             32.7%;   Score 318.5;   DB 10;   Length 528;
Best Local Similarity    43.3%;   Pred. No. 7.7e-22;
Matches       81; Conservative 22; Mismatches 69; Indels 15; Gaps 4;

QY      6 QYPINFTTAGATVQSNTNFIRAVGRILT-----VLNPNRVGLPINORFILVELSNHA 57
Ddb        |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
QY      1 QDOVKIFTEGATSQSYQFTEALQRLLTGGLIHDPVLPDPTTBERRIYITVELSNSE 60
Ddb         |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
QY     58 ELSVTLALDVTVNAVYVGVTRAGNSAYFFHPDNQEDAEAITHLFTDVQNRYTFAGSNDRI 117
Ddb         |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
QY     61 RESIEVGIDTVNAVVAAGSQSFV---RDAPASATYLFPFGTO-RVSLRFDGSIGDL 116
Ddb         |:|:|:|:|:|:|:|:|:|:~

```



OX NCB1\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RA Van Damme E.J.M.;  
 RT "Characterization and cloning of lectins and ribosome-inactivating  
 RL proteins from *Sambucus nigra* leaves.";   
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOLYOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AF409135; AAL04123.1; -  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS50231; RICIN\_B\_LCTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolase; Toxin.  
 SQ SEQUENCE 563 AA; 62242 MW; 07F7CBEDCF33BF10 CRC64;

Query Match 31.5%; Score 307.5; DB 10; Length 563;  
 Best Local Similarity 38.4%; Pred. No. 9.5e-21;  
 Matches 73; Conservative 36; Mismatches 56; Indels 25; Gaps 5;

QY 7 YPIINFATTAGATVQSYTNFIRAVR-----GRLTVLPNRVGLPINQRFILVELSNH 56  
 Db ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||  
 28 YPSVSENLGAKSATYRDFLSNLRKTVATGTYEVNGLPVLRRSEVQVKSRFVLVPLTNY 87  
 QY 57 AELSVTIALDVTNAYVGYRAGNSAYFFHFDNODAEAL--THLFTDVQNRNYTFAPGGNY 114  
 Db ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||  
 88 NGNTVTILAVDVTNLYVYAFSGNANSYFF-----KDATEVQKSNLFVGTQKN-TLSFTGNY 141  
 QY 115 DRLEQLAGNLRENIELNGPFLBEAISALYYSTGGTQLPTLARSFICICIMISEAARFOY 174  
 Db ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||  
 142 DNLETAANTERESIELGPSPLDGAITSLYHGD-----SVARSLVVIQWSEAAARFY 194  
 QY 175 IEGEMTRIR 184  
 Db ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||  
 195 IEQEVRRSLQ 204

RESULT 15  
 Q8GT32 PRELIMINARY; PET; 563 AA.  
 AC Q8GT32;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein nigrin 1 precursor  
 DE (EC 3.2.2.22).  
 OS *Sambucus nigra* (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCB1\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RA Girbes T., Arias F.J., Antolin P.;  
 RT "Characterization and molecular cloning of Nigrin 1, a type two  
 RL ribosome-inactivating protein from leaves of elder (*Sambucus nigra*).";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF249280; AAN86130.1; -  
 KW Hydrolase; Glycosidase.  
 SQ SEQUENCE 563 AA; 62173 MW; 0EB236421FC5E04F CRC64;

Query Match 31.5%; Score 307.5; DB 10; Length 563;  
 Best Local Similarity 38.4%; Pred. No. 9.5e-21;  
 Matches 73; Conservative 36; Mismatches 56; Indels 25; Gaps 5;

QY 7 YPIINFATTAGATVQSYTNFIRAVR-----GRLTVLPNRVGLPINQRFILVELSNH 56  
 Db ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||  
 28 YPSVSENLGAKSATYRDFLSNLRKTVATGTYEVNGLPVLRRSEVQVKSRFVLVPLTNY 87  
 QY 57 AELSVTIALDVTNAYVGYRAGNSAYFFHFDNODAEAL--THLFTDVQNRNYTFAPGGNY 114  
 Db ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||  
 88 NGNTVTILAVDVTNLYVYAFSGNANSYFF-----KDATEVQKSNLFVGTQKN-TLSFTGNY 141  
 QY 115 DRLEQLAGNLRENIELNGPFLBEAISALYYSTGGTQLPTLARSFICICIMISEAARFOY 174  
 Db ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||  
 142 DNLETAANTERESIELGPSPLDGAITSLYHGD-----SVARSLVVIQWSEAAARFY 194  
 QY 175 IEGEMTRIR 184  
 Db ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||  
 195 IEQEVRRSLQ 204

Search completed: February 10, 2004, 16:26:40  
 Job time : 26.3521 secs